Evidential weight and legal admissibility of information stored electronically

Code of practice for the implementation of BS 10008

Fifth Edition

Peter Howes and Alan Shipman

Evidential weight and legal admissibility of information stored electronically

Evidential weight and legal admissibility of information stored electronically

Code of practice for the implementation of BS 10008

Peter Howes and Alan Shipman

bsi.

First published 1996
Second edition 1999
Third edition 2004
Fourth edition 2008
Fifth edition 2014

by

BSI Standards Limited
389 Chiswick High Road
London W4 4AL

© The British Standards Institution 2014

All rights reserved. Except as permitted under the Copyright, Designs and Patents Act 1988, no part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means – electronic, photocopying, recording or otherwise – without prior permission in writing from the publisher.

Whilst every care has been taken in developing and compiling this publication, BSI accepts no liability for any loss or damage caused, arising directly or indirectly in connection with reliance on its contents except to the extent that such liability may not be excluded in law.

While every effort has been made to trace all copyright holders, anyone claiming copyright should get in touch with the BSI at the above address.

BSI has no responsibility for the persistence or accuracy of URLs for external or third-party internet websites referred to in this book, and does not guarantee that any content on such websites is, or will remain, accurate or appropriate.

The rights of Peter Howes and Alan Shipman to be identified as the authors of this Work have been asserted by them in accordance with sections 77 and 78 of the Copyright, Designs and Patents Act 1988.

Typeset in Great Britain by Letterpart Limited, www.letterpart.com

Printed in Great Britain by Berforts Group, www.berforts.co.uk

British Library Cataloguing in Publication Data

A catalogue record for this book is available from the British Library

ISBN 978 0 580 85676 1

Contents

Foreword

Evidential weight and legal admissibility of information stored electronically — Code of practice for the implementation of BS 10008 (referred to in this document as 'the Code') is primarily concerned with the authenticity, integrity and availability of electronically stored information, to the demonstrable levels of certainty required by an organization. It is particularly applicable where this stored information may be used as evidence in disputes inside and outside the legal system.

This is the fifth edition of the Code, which was first published by BSI in 1996. This edition is an editorial revision of the fourth edition (BIP 0008-1 (2008)). It is technically similar, with an extension of its scope to include information stored in databases and other electronic systems. It has also been restructured in recognition of the publication of BS 10008:2014, *Evidential weight and legal admissibility of electronic information — Specification* and can be considered to be a guide to the implementation of the British Standard in relation to information stored electronically.

Users of the previous editions should consider the advantages of assessing their information management systems in the light of this new edition, and amend their systems and/or documentation where appropriate. Guidance is given in Annex A of this part of the Code on the major differences between this version and the previous versions.

This publication is the first part of BIP 0008, which is made up of the following:

- BIP 0008-2 (2014), *Evidential weight and legal admissibility of information transferred electronically — Code of practice for the implementation of BS 10008*;
- BIP 0008-3 (2014), *Evidential weight and legal admissibility of linking electronic identity to information — Code of practice for the implementation of BS 10008*.

The Code is published by BSI in recognition of the large number of implementations of electronic information management systems, and of the continuing uncertainty about the legal acceptability of information stored on these systems. It provides good practice guidance for the electronic creation, storage and retrieval of information.

Acknowledgements

The Editors would especially like to thank the BSI Legal Admissibility Editorial Board and Panel and committees IDT/1, *Document management applications* and IDT/1/-/5, *Revisions of BS 10008* for their contribution to the current and previous editions of this publication, in particular for their business foresight and tireless reading of the manuscript. Their suggestions for improvements added value to the final publications.

The members of IDT/1 are Martin Bailey, Ian Curington, Aandi Inston, Marc Fresko, Peter Howes, Philip Jones, Andrew Kenny, Bill Mayon-White, Roger S Poole, Nick Pope, Ian Walden, Leonie Watson, Andrew Pibworth, Neil Pitman, Alan Shipman and Tom Wilson.

The members of IDT/1/-/5 are Elisabeth Belisle, Bernie Dyer, Peter Howes, Richard Jeffrey-Cook, Bill Mayon-White, Roger S Poole, Alan Shipman, Rod Stone and Tom Wilson

In particular, we would like to thank Jennifer Carruth from BSI for her excellent advice and copy-editing work on BS 10008:2014.

Peter Howes
Alan Shipman
(Editors)
Group 5 Training Limited

The Editors would also like to thank the following organizations for reviewing the previous editions of this publication:

Association of Chief Police Officers (ACPO);
Association for Payment Clearing Services (APACS);
British Computer Society (BCS) – Information Risk Management & Audit (IRMA) specialist group;
National Audit Office (NAO);
Police Information Technology Organisation (PITO);
The National Archives (TNA).

The first version of the Code, published in 1996, was authored by the following: Rob Allen, Bernard Dyer, Ian Galbraith, Bill Mayon-White, Roger Peggram, Alan Shipman (Editor) and Malcolm Smith.

The first version of the Code was sponsored by the leading information management users and trade organizations in the UK, namely: Computing Suppliers Federation, Information and Document Management Association, Legal Images Initiative consortium and UKAIIM Standards Committee.

The first version of the Code was developed with the assistance of a group of leading UK companies, consultants and associations: Advent Imaging Limited, AMEC Construction, APACS, Ashford Borough Council, Association of Computer Telephone Integration Users and Suppliers (ACTIUS), Autotrol Technology Limited, Bell and Howell, Bird and Bird, Blueprint, Brighton University, British Computer Society Computer Audit Specialist Group (BCS-CASG), British Standards Institution (BSI/DISC), Centre for Commercial Law Studies (University of London), Cheque and Credit Clearing Company Limited, Cimtech Limited, Concordium Software Limited, Document Image Technology, European Commission, European Electronic Messaging Association (EEMA), Fujitsu Europe Limited, Group 5 Training Limited, Headway Technology Group, Hewlett-Packard Limited, Imtec Group, Intergraph (UK) Limited, Kodak, Lloyds TSB Group, Lombard Document Systems Limited, London Borough of Enfield, London Transport, Marc Fresko Consultancy, Maxoptix Europe Limited, Microgen (UK) Limited, MR Group plc, National Westminster Bank, North Hampshire Hospital, Oki Systems (UK) Limited, OMS Services Limited, Q Star Limited, Registration Board for Assessors, Royal Bank of Scotland, SBC Warburg, Scottish Nuclear Limited, SNS Systems Inc., Society of Archivists, Sony (UK) Limited, SSI Microcad, Staffware plc, Tekdata

Limited, Tower Technology, Trimco Enterprises Limited, UK Banks Credit Card Committee, Wicks and Wilson Limited, Workflow & Groupware Strategies, Workflow Management Coalition (WfMC), Xerox Imaging Systems and 3M (UK) plc.

BSI would also like to thank the following who reviewed the fifth edition of this book:

John Avallanet, Managing Director & Principal, Cerulean Associates LLC;
Diane Shillito, Quality Systems Manager, CDS;
Neil Maude, General Manager, Arena Group;
Elisabeth Belisle, Managing Director, Scandox.

Introduction

Management summary

It is essential that organizations are aware of the value of the information that they store, and that they execute their responsibilities under the 'duty of care' principle. This Code of Practice gives detailed guidance on the issues of information management, information security management and legal/regulatory requirements. The Code is arranged based on BS 10008, *Evidential weight and legal admissibility of electronic information — Specification*, and can thus be used as a guide to the implementation of this British standard.

Information security is significant when discussing legal admissibility issues. Where legal admissibility is being assessed, the main discussion is likely to be related to the authenticity of the stored information. When the electronic information was captured by the storage system, was the process secure? Was the correct information captured, and was it complete and accurate? During storage, was the information changed in any way, either accidentally or maliciously? When responding to these questions, information security implementation and monitoring will be significant evidence when asked to demonstrate authenticity.

Information as an asset

The board of directors (or other equivalent group) of any organization is responsible for the conduct of that organization in every way – financially, operationally, legally and ethically. Specifically, it has responsibility for its assets and their use. Many responsibilities of a board of directors concern the activities and processes of the organization, for example investment for a new product, selling in a new market or building a new plant. But some of the most important responsibilities are defined functionally by subject, for example financial affairs and human resources. One such subject is information – not information systems but the stored information itself.[1]

Organizations operate by producing, transmitting and digesting information. The right information at the right place at the right time is essential for effective conduct of business. Equally, the misuse, copying, theft, loss and abuse of information can be, and has very publicly been, the cause of scandals and business failures.

Information is required in every activity and every function, thus proper control of information and care in its use has always been a subject of concern. Modern computers and communications systems can store information, process it and make it accessible in ways never before achieved. This can be of great additional benefit to business but can also enhance opportunities for misuse, theft, loss and abuse and, in particular, indiscriminate dissemination of information.

In some organizations, it is accepted that some types of business information are assets, for example, customer and services information and intellectual property such as patents and copyright. All information in an organization, regardless of its purpose, should be properly identified, even if identification is not required for accounting purposes, for consideration as an asset of the business. On the other hand, the retention of information past its retention period can also be a business liability, for example an increased cost of storage.

Most organizations have extensive experience in the subjects and functions they address. Relatively few organizations have experience in the acquisition, processing, storage and transmission of information and fewer still in the responsibilities that arise when information is considered a business asset.

[1] In 1995, with the support of the KPMG IMPACT team, the Hawley Committee produced a report, '*Information as an Asset*', with an objective of creating a set of guidelines for boards of directors on policies and procedures for managing information.

Purpose of the Code

Users of electronic information management systems are being asked by their companies, government departments and other employers to commit key records and documents under their control to electronic media. The application of these systems is changing the way in which many aspects of business and organizational life are operated, and is creating an electronic legacy for their successors, as paper documents are increasingly replaced by many forms of electronic information storage. Different electronic storage systems and devices have their own inherent advantages and limitations and existing systems will, at some later stage, be replaced or become obsolete. The purpose of the Code is to assist organizations in dealing with the implications, specifically concerning evidential and legal issues, of this technological evolution.

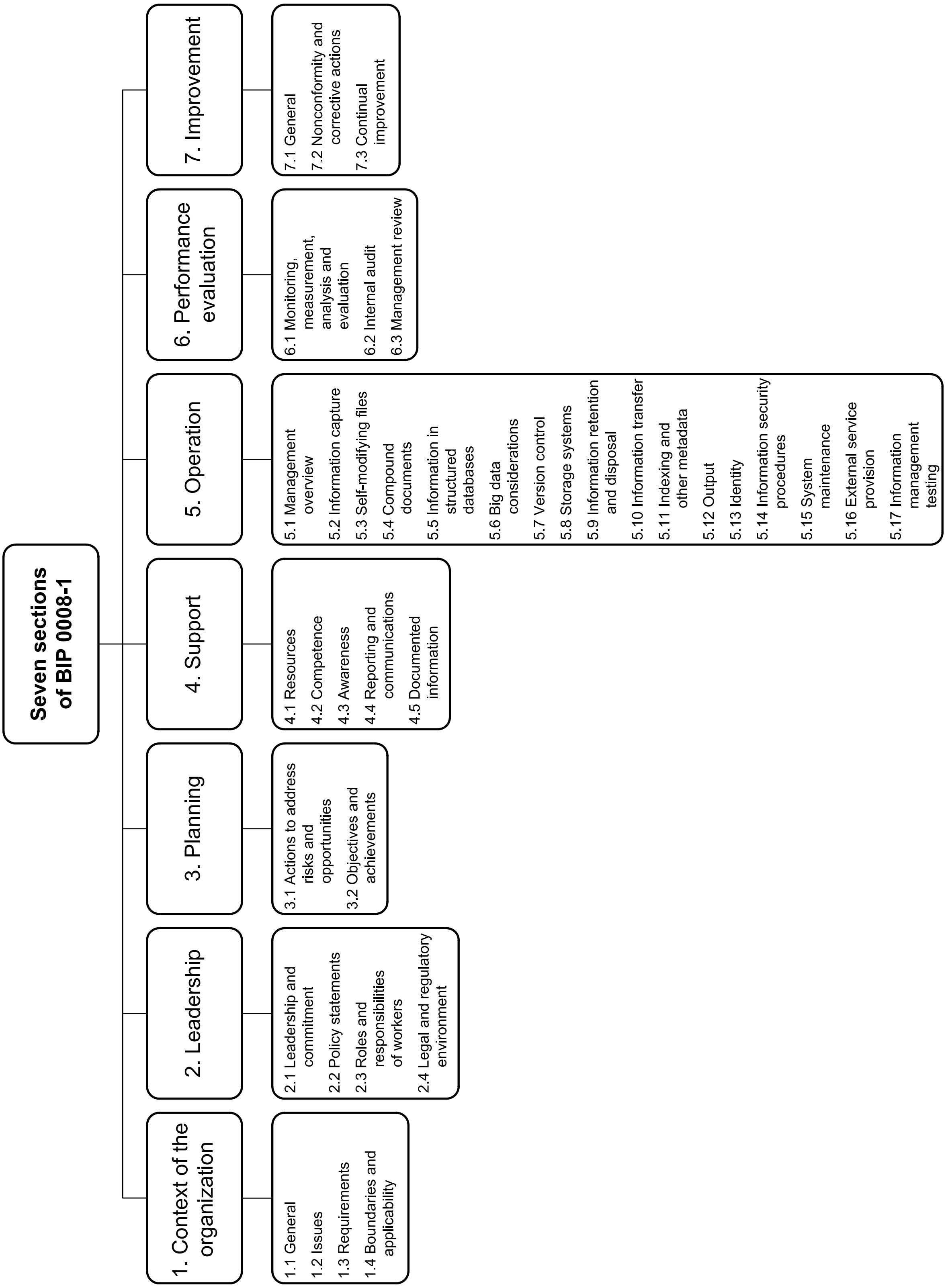
The Code provides a framework and guidelines, based on the provisions of BS 10008 that identify key areas of good practice for the implementation and operation of such electronic storage systems, whether or not any information held therein is ever required as evidence in the event of a dispute. As such, compliance with the Code (and therefore with BS 10008) should be regarded as a demonstration of responsible business management.

A more detailed explanation is provided in Annex B, to assist readers new to the subject or seeking background rationale for the guidance to good practice in the rest of the document.

Management framework

Chapters 1 to 7 of the Code are structured along the lines of the standardized structure of the ISO Management System Standards, such that its implementation can be synchronized with other management systems such as BS ISO/IEC 27001:2013 *Information technology — Security techniques — Information security management systems — Requirements* where appropriate.

Previous versions of the Code were structured according to the four phases of the Plan-Do-Check-Act, or PDCA for short, which had been adopted by the majority of management systems standards. Recent changes to the structure of these management system standards (such as those for quality management (BS EN ISO 9001:2008, *Quality management systems — Requirements*), environmental management (BS EN ISO 14001:2004, *Environmental management systems — Requirements with guidance for use*) and information security management (BS ISO/IEC 27001)) have taken place. The PDCA model emphasizes how critical it is for effective risk management that an ongoing management process is in place, and is exercised.


Seven sections of BIP 0008-1
1. Context of the organization
1.1 General
1.2 Issues
1.3 Requirements
1.4 Boundaries and applicability
2. Leadership
2.1 Leadership and commitment
2.2 Policy statements
2.3 Roles and responsibilities of workers
2.4 Legal and regulatory environment
3. Planning
3.1 Actions to address risks and opportunities
3.2 Objectives and achievements
4. Support
4.1 Resources
4.2 Competence
4.3 Awareness
4.4 Reporting and communications
4.5 Documented information
5. Operation
5.1 Management overview
5.2 Information capture
5.3 Self-modifying files
5.4 Compound documents
5.5 Information in structured databases
5.6 Big data considerations
5.7 Version control
5.8 Storage systems
5.9 Information retention and disposal
5.10 Information transfer
5.11 Indexing and other metadata
5.12 Output
5.13 Identity
5.14 Information security procedures
5.15 System maintenance
5.16 External service provision
5.17 Information management testing
6. Performance evaluation
6.1 Monitoring, measurement, analysis and evaluation
6.2 Internal audit
6.3 Management review
7. Improvement
7.1 General
7.2 Nonconformity and corrective actions
7.3 Continual improvement

General

Scope

The Code describes the implementation and operation of information management systems that store information electronically and where the issues of authenticity, integrity and availability as required by legal admissibility and evidential weight are important.

DEFINITIONS (see also Annex C)

Authenticity – trustworthiness of origin and evidential content

Integrity – retention of the evidential content of the information

Availability – accessibility of the information as required

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

The Code is for use with any information management system that stores information electronically, using any type of electronic storage medium including write-once-read-many (WORM) and rewritable technologies.

The Code is also for use with any type of database or other electronic system. Database files may potentially contain any type of data: for example, coded-character data, formatted text, images, computer-aided design (CAD) drawings, moving and still video images or voice data, or any combination of these. Database files may contain data of more than one type, and/or of more than one image. Database files may also include internally generated files, such as log files and audit trails.

Database files may be created by the information management system itself or by its users, or they may be imported into the system. The Code covers all such database files, whether created and/or imported directly or through a network, from the time at which the system assumes control of the database file.

The Code does not cover processes used to evaluate the authenticity of information prior to it being imported into the system. However, it can be used to demonstrate that output from the information management system is a true record of what was imported.

The Code is also for use in the identification and development of policies and procedures as specified in BS 10008, in relation to the storage of electronic information. A companion to the British Standard and to the Code is the compliance workbook, BIP 0009 (2014), *Evidential weight and legal admissibility of electronic information — Compliance workbook for use with BS 10008*. This workbook enables a comprehensive assessment to be made of the user's information management system for compliance with the British standard. Completion of a copy of BIP 0009 for each system and associated storage and retrieval processes provides one means of satisfying key elements of the audit trail.

Voice, audio and video data

Data files may contain voice, audio and/or video information. Such files can be managed in accordance with the Code.

For all such files, once the recording is frozen, the file needs to be treated in the same way as any other data file as far as the Code is concerned.

Where the recording of voice, audio and/or video data is not under the control of the information management system, the recording system needs to have control of file integrity that is at least as good as that imposed by the Code for other types of information capture.

Where voice, audio and/or video data is stored, procedures for authentication of the source of the data need to be documented.

Applicability

The Code is applicable to any system that stores information electronically. It covers aspects of the information management processes that affect the use of information in normal business transactions, even where legal admissibility per se is not an issue. Such aspects include the legibility, accuracy and completeness of the stored information, and the transfer of the information to other systems.

Technology

It is important to utilize reliable and trustworthy technology to store electronic information over a long period of time, potentially with the implementation of replacement technologies. Each part of the system needs to be chosen with care, taking into account the possible need to demonstrate 'proper' and 'appropriate' working of the system sometime in the future. This demonstration may need to encompass both the technology itself and the methods by which it was configured and used. The technology sections cover particular aspects of technology (e.g. the storage media used) as well as critical aspects of configuration (e.g. how access to the system is managed).

The users

The Code is intended for:

- end user organizations that wish to ensure that information created by, entered into and/or stored within their information management systems may be used with confidence as evidence in any dispute, within or outside a court of law;
- integrators and developers of information management systems that provide facilities to meet user requirements.

Objectives

The objectives of the Code are to:

- improve reliability of, and confidence in, stored information;
- maximize the evidential weight that a court or other body may assign to presented information;
- provide confidence in inter-company trading;
- provide confidence to external inspectors (e.g. regulators and auditors) that the organization's information and business practices are robust and reliable.

The Code may be used as a common reference for business activities within and between organizations and for subcontracting or procurement of IT services or products.

Compliance

Each chapter of the Code contains a general description of the issues being addressed, followed by a list of 'key issues'. These key issues indicate the critical compliance points that need to be taken into consideration, and acted upon where appropriate, before compliance with the recommendations of the Code can be claimed. Compliance is claimed on a voluntary basis, by self-certification.

A compliance workbook (BIP 0009) has been published by BSI to enable an assessment of compliance with BS 10008 to be demonstrated. Where critical compliance points from the Code are not specifically included in the British standard, these points are included as an optional component in the compliance workbook.

Where compliance on a self-assessment basis is claimed, recommended compliance statements as shown in 6.3.4 should be used. See also 6.3 for further information on compliance audits.

It should be noted that, where compliance is assessed by a third party, liability for compliance will normally remain the responsibility of the organization.

1 Context of the organization

1.1 General

This section of the Code relates to Clause 4 of BS 10008, 'Context of the organization'.

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.

In order to ensure that this information is well managed, and to meet its business needs, the organization needs to define and implement good management practices. Information, like any other asset, needs to be classified, structured, validated, valued, secured, monitored, measured, managed and disposed of when appropriate, efficiently and effectively.

1.2 Issues

The organization needs to determine the external and internal issues that are relevant to its purpose and that may affect the authenticity and integrity of the information that it uses.

Typical issues that may be relevant include:

a) the size and complexity of the organization;
b) the level of business risk attached to being unable to demonstrate authenticity and integrity of stored information;
c) drivers for business efficiency improvements;
d) specific stakeholder requirements;
e) the existing technology and infrastructure systems.

Policy statements as described in 2.2 should take into account those issues that are agreed to be relevant to the ability to demonstrate authenticity and integrity of information stored electronically.

When reviewing the relevant issues, a risk management process is the most appropriate to use when deciding upon actions to be undertaken. BS ISO 31000:2009, *Risk management — Principles and guidelines* provides principles, framework and a process for managing risk. It can be used by any organization regardless of its size, activity or sector. Using BS ISO 31000 can help organizations increase the likelihood of achieving objectives, improve the identification of opportunities and threats and effectively allocate and use resources for risk treatment.

1.3 Requirements

When establishing or reviewing the systems and/or processes that manage the evidential weight of information stored electronically, the organization needs to determine:

a) stakeholders that are relevant to the authenticity and integrity of information;
b) the requirements of these stakeholders relevant to that information;
c) the requirements for information stewardship within the organization.

NOTE: The requirements of stakeholders may include legal and regulatory requirements and contractual obligations.

Typical stakeholders may include:

- owners, managers and staff of the organization;
- third parties with contracts or similar agreements with the organization;
- clients and customers in receipt of services provided by the organization;
- the public where public services are involved;
- regulatory bodies;
- government bodies;
- external audit bodies;
- legal advisers.

The requirements of each stakeholder need to be taken into consideration when producing policy statements (see 2.2).

Information stewardship should be managed by the identification of Information Asset Owners (IAOs) who will typically be those responsible for the processes that generate the information asset in question.

1.4 Boundaries and applicability

The organization needs to determine the boundaries and applicability of the authenticity and integrity of the information it uses in order to establish its scope.

When determining this scope, the organization needs to consider:

a) the external and internal issues referred to in 1.2;
b) the requirements referred to in 1.3; and
c) interfaces and dependencies between activities performed by the organization, and those that are performed by other organizations.

The scope needs to be available as part of the policy document.

In many organizations, the authenticity and integrity of information will only be of importance to part of the overall information asset. As part of a project to implement BS 10008 and the Code, individual information assets need to be identified and a decision taken as to whether each should be included within the scope of the related policy statement.

2 Leadership

2.1 Leadership and commitment

This section of the Code relates to Clause 5 of BS 10008, ‘Leadership’.

Top management needs to demonstrate leadership and commitment with respect to the management of the information authenticity and integrity by:

a) ensuring the information management policies and objectives are established and are compatible with the strategic direction of the organization;
b) ensuring the integration of the information management system requirements into the organization’s processes;
c) ensuring that the resources needed for the information management system are available;
d) communicating the importance of effective information management and of conforming to the information management system requirements;
e) ensuring that the information management system achieves its intended outcome(s);
f) directing and supporting persons to contribute to the effectiveness of the information management system;
g) promoting continual improvement; and
h) supporting other relevant management roles to demonstrate their leadership as it applies to their areas of responsibility.

2.2 Policy statements

2.2.1 General

BS 10008 specifies the contents of an electronic storage policy statement, which covers the scope of the policy, requirements for procedures and technology and the responsibilities of the management of the systems. The standard also specifies the requirement for the top management team to set a clear policy direction and demonstrate support for, and commitment to, the management of the electronic information through the issue and maintenance of an information management policy.

The standard also specifies the contents of an information security policy within which the information management policy operates.

Policy documentation needs to be retained in compliance with the retention schedule.

KEY ISSUE

> Retain approved policy documents in line with the retention schedule.

2.2.2 Information management policy statement

2.2.2.1 Structure

This section describes documentation that states the organization’s information management policy. Availability of this documentation should, when combined with appropriate proof of compliance, demonstrate (e.g. to a court of law) that responsible information management is part of the organization’s normal business practice.

An information management policy statement (the ‘policy statement’) should be produced, documenting the organization’s policy on information management and storage, as applicable to the information management system. Policy objectives should be clearly stated, such that they can be easily interpreted and implemented.

NOTE: A fundamental recommendation of the Code is the formalization, agreement and implementation of a retention schedule for stored information. Where reference is made to the policy document in the rest of the Code, the retention schedule is included in such a reference.

The policy statement should be approved by the top management of the organization, and should be reviewed for relevance and content at regular intervals. The frequency for review should be appropriate to the application. This period will typically be the same as the normal procedural audit cycle within the organization (e.g. annual or in the event of major changes to the system).

Annex D includes an example policy statement, which may be used during the drafting of an organization’s policy statement. It contains some ‘typical’ statements that may be appropriate in many policy statements.

The policy statement should contain, as a minimum, the following. Other sections may be added where appropriate.

Topic	Content	Section of BIP 0008-1
Scope	Specifies which information ‘types’ are covered by the policy	2.2.2.2
Information classification	States the policy regarding information classification (where used)	2.2.2.3
Standards related to information management	States the policy regarding relevant information management standards	2.2.2.4
Retention and disposal schedules	Defines retention periods and disposal policies	2.2.2.4.3.2
Legal and regulatory consultations	Defines the requirements for consultations with appropriate legal and/or regulatory bodies, and with others where necessary	2.2.2.5
Roles and responsibilities	Defines roles and responsibilities for information management functions and for compliance	2.2.2.6
Storage technology	States the policy regarding the type of technology to be used for storage	2.2.2.7
Electronic file formats	States the policy regarding electronic information formats and version control; this includes electronic file formats and database management systems and database schemas	2.2.2.8
Auditing	Defines requirements for auditing relative to particular document types	2.2.2.9

Table 1 – Information management policy topics

KEY ISSUES

> Develop an information management policy statement and have it approved by top management.

> Ensure it is reviewed at regular intervals, as appropriate to the application.

2.2.2.2 Scope

The policy statements described in 2.2.2.1 may not cover all the different types of information that the organization uses. The information to be included in its scope should be identified and grouped into types, with the policy for all information within a type being consistent. Where a retention and disposal schedule (see 2.2.2.4.3.2) exists, it may be appropriate to use the same information type groups.

NOTE: Where formal records management procedures are implemented, information types will generally be used in classification schemes (file plans) and/or taxonomies. Where this is the case, the same file grouping and naming conventions should be used.

EXAMPLES of information types

Information types specified by reference to application (e.g. ‘financial projections’, ‘invoices’ or ‘customer address list’)

Information types specified by reference to generic group (e.g. ‘accounting data’, ‘customer documents’ or ‘manufacturing documents’)

The policy statement should include, as one or more information types, all documents produced in conformance to BS 10008, as well as system and application audit trails (see 4.5.3).

KEY ISSUE

> All information types should be included in the policy statement.

An alternative to this approach of including only identified information types in the policy scope can be adopted, since there is a risk that the organization may need to produce absolutely any of its information assets in the event of a dispute. The key consideration is whether the organization should have any information that is totally unmanaged, even the information that is regarded as of very low value to the organization.

In this case, it is worth considering whether the policy scope should include a ’default’ information type (with associated policy and policy implementation and compliance requirements) as well as the explicitly identified information types. Information outside the scope of the explicitly identified types should be managed according to the default.

2.2.2.3 Information classification

In some applications, it is appropriate to implement an information classification system, typically used to indicate the accessibility of particular documents to workers and other individuals. In government and other public bodies, this is often indicated by the use of security ‘labels’ such as ‘top secret’, ‘classified’ or ‘publicly available’. In the private sector, security classification schemes may be aligned to departmental requirements (such as accounts, credit control or customer services).

It is important to ensure that the information classification system is simple to understand, and thus easy to implement. It should be based on need, priority and degree of protection needed. Excessive classification levels should be avoided.

Where an information classification system is implemented, the policy statement should include with each information type the details of its classification.

KEY ISSUES

> Include in the policy statement against each information type its information classification (where used).

> Keep the information classification system simple and understandable to all users.[2]

2.2.2.4 Standards related to information management

2.2.2.4.1 General

Frequently, business benefits can be achieved by complying with relevant national or international standards, codes of practice or other guides. The policy document should state whether all or specific parts of any such publication(s) are to be complied with.

KEY ISSUE

> Where business benefits can be derived from compliance, the appropriate standard(s) (and other guides) should be implemented.

2.2.2.4.2 Quality management

Where the organization operates a quality management system (such as one in accordance with BS EN ISO 9001, *Quality management systems — Requirements*), whose scope includes part or all of the information management system within the scope of compliance, then it is recommended that all documentation that BS 10008 requires should be included in the quality management system.

KEY ISSUE

> Where a BS EN ISO 9001 or other quality management system is implemented, it is recommended that all documentation and procedures required by BS 10008 should be included within its scope.

2.2.2.4.3 Records management

2.2.2.4.3.1 General

A British Standard on the subject of records management was originally published in 2001. This is BS ISO 15489-1:2001, *Information and documentation — Records management — Part 1: General*, and includes an international standard and a related technical report. At the time of publication of this Code, BS ISO 15489-1 is undergoing a significant revision. Users should ensure that the latest version of BS ISO 15489-1 is used.

Many of the documents stored within the information management system will be records, and thus should be managed in accordance with the standard. Annex E contains a cross-reference listing,

[2] One such scheme used within government is the Government Security Classifications. For information, see www.gov.uk/government/publications/government-security-classifications

detailing the various requirements of BS ISO 15489-1 and their mapping to the Code. This will allow users of the Code to implement its recommendations within the requirements of the records management regime.

> **DEFINITIONS from BS ISO 15489-1:2001**
>
> Records – 'information created, received and maintained as evidence and information by an organization or person, in pursuance of legal obligations or in the transaction of business'
>
> Records management – 'field of management responsible for the efficient and systematic control of the creation, receipt, maintenance, use and disposition of records, including processes for capturing and maintaining evidence of and information about business activities and transactions in the form of records'

2.2.2.4.3.2 Retention and disposal schedules

A retention schedule should be established for each information type included in the policy statement and agreed with all relevant departments and personnel within the organization. They should be reviewed at regular intervals. Where appropriate, legal advice may be required to ensure that minimum legal and/or regulatory retention periods, as well as business requirements, are complied with. For further information and guidance on retention schedules, see:

- The National Archives (previously the Public Record Office);[3]
- The Information and Records Management Society.[4]

All documentation that BS 10008 and the Code require to be kept should be covered by the retention schedule.

There should also be a disposal schedule, stating the organization's policy for the controlled destruction (or other process such as archival preservation) of information at the end of its retention period. In many instances, the disposal process will involve the physical destruction of the information (including all copies). In other cases, information may be passed to archives (to allow it to be used for research and/or statistical purposes), in which case policies are needed that detail which documents (and maybe which information in these documents) are to be archived.

These schedules should be included in, or referenced by, the policy statement.

KEY ISSUE

> \> A retention schedule and a disposal schedule should be established and included in, or referenced by, the policy statement, covering all information types included in the policy statement.

2.2.2.5 Legal and regulatory consultations

2.2.2.5.1 Consultations

It is important to be aware of the various legal and regulatory requirements that the organization needs to comply with. Some (or maybe all) of these require the keeping of information. Some legal advisers or regulators may be unfamiliar with electronic information management systems, and so will need to seek consultation before a 'mainly electronic' process is implemented.

Typical bodies that may need to be consulted are:

[3] www.nationalarchives.gov.uk/information-management/manage-information/

[4] www.irms.org.uk

- regulatory bodies;
- government bodies;
- external audit bodies;
- legal advisers.

These consultations may need to include the following topics:

- legal issues (civil law and/or company law);
- government regulations;
- special regulations (applicable to particular sectors).

These consultations may be appropriate at different levels, which may include:

- national and international law;
- industry sector regulation;
- community by-laws;
- organizational policies;
- departmental procedures;
- individuals' rights.

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

KEY ISSUES

> Relevant consultations should be made to ensure that the organization is aware of, and thus can comply with, all appropriate legal and/or regulatory requirements.

> The results of all consultations should be documented in (or referenced by) the policy statement.

2.2.2.5.2 UK legal and regulatory bodies

When operating within the UK, particular bodies may need to be contacted as part of the organization's consultations. UK users of the Code may consider consulting the following organizations:

- Companies House;
- HM Revenue & Customs.

In specialized industry sectors, organizations may need to comply with the requirements of bodies such as:

- the Financial Conduct Authority (FCA);
- the Civil Aviation Authority;
- the Food Standards Agency (and possibly the US Food and Drug Administration);
- industry regulators.

Where relevant, it may also be helpful to consult organizations and individuals such as:

- the company secretary;
- the records manager;
- internal and external legal advisers;
- internal and external auditors;
- other government bodies, for example:
  - The National Archives;
  - the Office of Government Commerce;
  - the e-Government Unit;
  - the Health and Safety Executive;
  - the National Audit Office (for public organizations audited by the NAO);
  - the Information Commissioner's Office.

2.2.2.5.3 Paper replacement systems

Historically, the majority of information has been stored on paper. Organizations that carry out auditing activities, typically resulting from a legal or regulatory requirement, anticipate inspecting these paper records. Typical examples of such organizations are HM Revenue & Customs and the NAO.

Where a change to electronic working is planned, prior permission will generally be required from the appropriate authorities.

KEY ISSUE

> Ensure that prior permission to replace paper records with electronic records from all external bodies that carry out audit activities on the organization has been obtained.

2.2.2.6 Roles and responsibilities

There are a number of information management responsibilities that should be identified, to ensure compliance with BS 10008 and the Code. For further guidance, see 2.3.

2.2.2.7 Storage technology

Different types of storage technology have different long-term storage characteristics. Most organizations will store information on a variety of media: paper, microform and/or electronic (write-once and rewritable/erasable). In some applications, specific pieces of information may, throughout their retention periods, be stored on different media at different times, for example, where hierarchical storage management (HSM) systems are used (see 5.8).

In order to facilitate effective storage, the organization will need to have policies regarding the use of specific types of medium for different information storage requirements (e.g. access requirements, retention periods and security requirements). These policies should be used to decide the medium to be used for the storage of each information type. These policies should be included in the policy statement.

KEY ISSUE

> Include the approved storage technology medium for each information type in the policy statement.

2.2.2.8 Electronic file formats

2.2.2.8.1 General

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, data migration procedures can be implemented to ensure long-term retrieval of the stored information (see 5.8.7).

2.2.2.8.2 Electronic file formats

The policy statement should contain details of the approved electronic file formats that may be used for each information type. Where the electronic file may be subjected to compression techniques (see 5.10.3.2), their use should be documented.

KEY ISSUE

> Include the approved electronic file format and compression policies for each information type in the policy statement.

2.2.2.8.3 Database management systems and schemas

The policy statement should contain details of the approved database management systems (DBMS) and database schemas. Changes to approved DBMS and database schemas should be under an approved version control system (see 5.7.1.3).

In the event of the DBMS and schema for a particular application or service being unavailable or unpublished proprietary intellectual property of a third party, the organization will need to assess the implications of the unavailability of this information and, if deemed acceptable, document within the policy statement.

2.2.2.9 Auditing

Records should be kept of information management system historical activities, changes, configurations, events or decisions that may need to be accessed in the future, as additional evidence to support stored information.

Audit trails will need to contain sufficient and necessary information to provide evidence of the authenticity of stored information, including details of any changes to it. The content of the audit trails should be agreed with all relevant departments within the organization, including audit, compliance and legal, and should be documented or referenced in the policy statement.

KEY ISSUES

> Audit trails should contain sufficient information to be able to demonstrate all necessary historical activities relating to the system and to stored data.

> Details of audit trail requirements should be included in the policy statement.

2.2.3 Information security policy

2.2.3.1 Security elements

To fulfil the duty of care objective, the organization needs to action the following.

Topic	Content	Section of BIP 0008-1
Scope	Specifies the information security policy	2.2.3.3
Management objectives	States the objectives of the information security policy	2.2.3.4
Security classification	States security requirements for different information categories	2.2.3.5
Roles and responsibilities	Details roles and responsibilities in relation to information security management	2.2.3.6
Segregation of roles	Details requirements related to the segregation of roles	2.2.3.7
Access rights	Details policy on staff access to information and on sharing information with third parties	2.2.3.8
Security breach management	Details policy for dealing with actual or suspected security breaches	2.2.3.9
Information security management standards	States policy for compliance with information security standards (such as BS ISO/IEC 27001)	2.2.3.10

Table 2 – Information security topics

2.2.3.2 Information security management framework

Where an appropriate information security management framework does not already exist, one should be established. It should control the information security management system as implemented within the organization. The framework should have as its objectives:

- approval and review of the information security management policy;
- monitoring of threats to information security;
- monitoring and review of security breaches;
- approval of major initiatives to enhance information security management.

KEY ISSUE

> Plan and implement an information security management framework.

2.2.3.3 Scope

To document the organization's security requirements, the organization should adopt an information security management policy, covering all elements of the information management system.

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

The information security management policy document should contain, as a minimum:

- the scope of the information security policy;
- a statement of the management objectives in respect to security;
- specific policy statements;
- requirements for different information classification categories;
- the definition and allocation of information security responsibilities;
- a policy for dealing with breaches of security;
- a policy regarding compliance with relevant standards;
- a policy for updating the policy.

The information security management policy document should be approved by the organization's top management. The organization should then agree and document appropriate levels of security for managing its information, in compliance with its stated information security policy.

KEY ISSUES

> Develop, authorize and implement an information security management policy.

> Ensure that the policy's scope includes the information management and storage systems.

2.2.3.4 Management objectives

All information, irrespective of the storage technology, 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 (see 2.2.3.5) that are used.

Traditionally, information security is considered a matter of confidentiality, to ensure that information is not available for viewing 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 compliance requirements.

A key recommendation of BS 10008 and the Code is to protect the integrity of stored information. When developing security measures, it is necessary to compare the risk of integrity being compromised with the cost of the implementation of such measures. Security measures need to include backup and other copies of stored information, as their integrity is of importance in circumstances in which they have been used as replacements for 'live' data, or for use as 'legal' archives.

There is also the important issue of 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 such instances, issues such as indexing accuracy and business continuity planning are important.

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

A management statement should be produced, detailing the organization's strategy for information security management, upon which the information security policy (see 2.2.3) is based.

EXAMPLE

An example of a management statement in relation to information security management is as follows.

The management objectives of this information security policy are to:

- ensure that information is accurate, complete and available as required;
- demonstrate continuity of information management services;
- minimize business damage by preventing or minimizing the impact of information security incidents.

2.2.3.5 Security classification

As discussed in 2.2.2.3, an information classification system may be in use. Where this is the case, the information security policy should reference the classification system, and should align security measures to that system.

KEY ISSUE

> Where an information classification system is in use, the information security policy should be aligned to that system.

2.2.3.6 Roles and responsibilities

There are a number of information security responsibilities that should be identified to ensure compliance with BS 10008 and the Code. For further guidance, see 2.3.

2.2.3.7 Segregation of roles

The segregation 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, segregation 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 segregation of roles should be considered:

- input reconciliation;
- quality control;
- data entry;
- information disposal;
- information security.

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

In some instances, segregation of two roles may not be feasible, due for example to the lack of availability of alternate workers. A justification needs to be documented.

EXAMPLE

Where 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 systematic sampling.

KEY ISSUES

> Identify any procedure that is self-checking, and where there is a potential risk to the integrity of procedures.

> Modify procedures where necessary to ensure reliable systems.

2.2.3.8 Access rights

It is important to control access to electronic data files, by the implementation of an access control system. In an information management system, typical access rights levels may be as follows:

- system manager;
- system administrator;
- system maintenance;
- authors or originators;
- information storage and indexing;
- information retrieval.

The information security policy should state the requirement for an access level management system, based on job functions.

KEY ISSUE

> The information security policy should include a policy for the management of access to IT systems, based on job role.

2.2.3.9 Security breach management

From time to time, security breaches, whether actual or suspected, arise with information management systems, and these breaches require emergency procedures to be implemented in order to recover from them. 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 recovery operations, an agreed and approved information security incident management plan should be implemented.

In some instances, the security breach may not appear to have affected the stored information. However, checks should be made to demonstrate that the integrity of the information has not been compromised.

In order to ensure that recovery from a security breach that resulted in major equipment, environmental or personnel failure has been successful, procedures should be developed, tested, implemented and maintained. Such procedures should ensure that the integrity of stored information is not compromised during their implementation.

Where complete recovery from the results of a security breach cannot be assured, details of all actual or potential compromise to the stored information should be documented.

KEY ISSUE

> Security breaches should be managed in a controlled manner, using procedures that ensure (or highlight any actual or potential issues concerning) the maintenance of the integrity of stored information, during and after an incident.

2.2.3.10 Information security management standards

BS ISO/IEC 27002:2013, *Information technology — Security techniques — Code of practice for information security controls*, is the UK reference document for information security management. Proof of compliance with the recommendations of this standard when implemented within the boundaries covered by the Code may provide helpful, supporting evidence in court or other dispute. It will indicate that the organization has exercised its duty of care, and it will assist the adjudicator in assessing the authenticity and integrity of the information.

BS ISO/IEC 27001 is an auditable specification for use in the certification of an information security management system against the standard.

Compliance with the recommendations of BS ISO/IEC 27002 or certification against BS ISO/IEC 27001 should not be regarded as an alternative to compliance with the recommendations of the Code.

Compliance with the requirements of BS ISO/IEC 27001 is predicated upon the proven quality management approach of demonstrably meeting a stated requirement, as defined by the organization in question. As such, this enables BS ISO/IEC 27001 to be applied to organizations as varied as a government department, a high street bank and a consultancy employing one or two people. Clearly the risks associated with these businesses are different and the information security measures adopted will, therefore, also be different.

The Code works on a different premise. When information is used as evidence in the event of a dispute, the maximum weight of evidence is not affected by the size or shape of the organization and its own view of security risks. It frequently depends upon the opinion of an independent arbiter. That view may well be affected by the opposing party in the dispute attempting to discredit the evidential value. Therefore, the Code is based upon more prescriptive, less interpretative principles, and compliance with its recommendations is significantly more absolute than compliance with the requirements of BS ISO/IEC 27001.

BIP 0071–0074 give guidance on the implementation of BS ISO/IEC 27001:

- BIP 0071 (2013), *Guidelines on requirements and preparation for ISMS certification based on ISO/IEC 27001*;
- BIP 0072 (2013), *Are you ready for an ISMS audit based on ISO/IEC 27001?*
- BIP 0073 (2013), *Guide to the implementation and auditing of ISMS controls based on ISO/IEC 27001*;
- BIP 0074 (2006), *Measuring the effectiveness of your ISMS implementations based on ISO/IEC 27001*.

Where IT services are being subcontracted to another organization (or another company within the same group), it would additionally be appropriate to implement BIP 0005 (2011), *A manager's guide to service management*, which gives guidance on service management.

KEY ISSUE

> Information security standards provide auditable systems that enable organizations to demonstrate a duty of care in relation to stored information.

2.3 Roles and responsibilities of workers

Roles and responsibilities for information management and information security management should be defined in the relevant policy statements. These roles and responsibilities may extend outside the organization, where third parties have access to systems.

The responsibilities in relation to information management may be indicated with reference to a job function (e.g. the IT manager).

Responsibilities should be assigned for:

- policy statement content;
- policy statement approval (typically the senior board member);
- policy statement periodic review;
- information content of each information type defined;
- the information management system;
- ensuring continued compliance.

The responsibilities in relation to information security management may also be indicated with reference to a job function (e.g. the information security manager).

Responsibilities should be assigned for:

- security policy content;
- security policy approval (typically the senior board member);
- security policy periodic review;
- security classification for each information type defined;
- the information security management system;
- ensuring continued compliance.

KEY ISSUES

> Identify responsibilities for information management (by means of reference to the relevant job function) in the policy statement.

> Identify responsibilities for information security management (by means of reference to the relevant job function) in the information security policy.

2.4 Legal and regulatory environment

This topic is discussed in Annex H.

3 Planning

3.1 Actions to address risks and opportunities

3.1.1 General

This section of the Code relates to Clause 6 of BS 10008, ‘Planning’.

When planning for the management of the authenticity and integrity of information, the organization needs to consider the issues referred to in 1.2 and the requirements referred to in 1.3 and determine the risks and opportunities that need to be addressed to:

a) ensure the information management system can achieve its intended outcome(s);
b) prevent, or reduce, undesired effects; and
c) achieve continual improvement.

The organization also needs to plan:

a) actions to address these risks and opportunities; and
b) how to:
   1. integrate and implement the actions into its information management system processes; and
   2. evaluate the effectiveness of these actions.

3.1.2 Risk assessment

Information management procedures are often developed in an unstructured way, by reacting to user requirements, security incidents and/or to available computer software tools. This approach on its own can easily leave gaps in information management, which are only filled at some later date, typically after a security breach. A more structured approach is to review the information assets of the organization and assign risk factors (based on asset value, potential threats, system vulnerability and likelihood of attack), on the basis of which appropriate, cost-effective information management procedures can be identified. An essential part of information management is the implementation of an appropriate security policy, which should be produced and approved, based on the risk assessment, and against which security measures can be developed and implemented.

NOTE: A review of this type generally requires security expertise and a range of appropriate technical skills.

The organization should undertake an information security risk assessment along these lines, 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 medium being used (e.g. WORM or rewritable).

On the basis of the results of the risk assessment, existing security measures should be reviewed for effectiveness. Factors such as the balance between the cost of implementation and the security achieved need to be taken into consideration during the review process. Where the review indicates that changes to security measures are appropriate, an action plan should be drawn up with new or amended security measures prioritized for implementation.

KEY ISSUE

> Perform a risk assessment of existing security measures, and implement cost-effective technology and/or procedures to fill any gaps found.

The risk assessment will lead to the acquisition of information and the creation of risk reports. These reports, backed up by the information used to develop the conclusions and recommendations in the reports, may provide useful evidence in relation to information management decisions made by the business.

Consequently, it is thus important to retain information related to risk assessments in line with an information retention schedule.

BS ISO 31000 provides principles and generic guidelines on risk management. It can be used by any public, private or community enterprise, association, group or individual. It can be applied throughout the life of an organization, and to a wide range of activities, including strategies and decisions, operations, processes, functions, projects, products, services and assets. It can be applied to any type of risk, whatever its nature, whether having positive or negative consequences.

KEY ISSUE

> Retain records of risk assessment methods and results in line with the retention schedule.

3.1.3 Risk treatment

The results of the risk assessment should be used to guide and determine the appropriate management action and priorities for managing information risk and implementing controls selected to protect against those risks.

BS ISO/IEC 27005:2011, *Information technology — Security techniques — Information security risk management* provides information security risk management guidance, including advice on risk assessment, risk treatment, risk acceptance, risk communication, risk monitoring and risk review.

BS ISO/IEC 27005 describes the input to a risk treatment process as a list of identified risks, prioritized according to the organization’s risk evaluation criteria. Risk treatment includes the identification and implementation of controls to reduce, retain, avoid or share the identified risks.

Risk treatment can be implemented by one or more of the following non-exclusive processes:

- risk modification;
- risk retention;
- risk avoidance;
- risk sharing.

Risk modification involves the addition, removal or modification of existing controls such that the residual risks can be re-evaluated.

Risk retention is the process of retaining an identified risk without further action. This is acceptable where the identified risk is within the agreed risk criteria.

Risk avoidance involves the removal of processes related to the risk, such that the risk is no longer present. This may be used where the cost of other forms of risk treatment are too costly to implement.

Risk sharing involves the sharing of the identified risks with other parties, such as by insurance or by subcontracting particular processes.

3.2 Objectives and achievements

The organization needs to establish information management objectives at relevant functions and levels.

The information management objectives need to:

a) be consistent with the information management policy;
b) be measurable (if practicable);
c) take into account applicable information management requirements, and results from risk assessment and risk treatment;
d) be communicated; and
e) be updated as appropriate.

The organization should retain information on the information management objectives.

When planning how to achieve its information management objectives, the organization needs to determine:

a) what will be done;
b) what resources will be required;
c) who will be responsible;
d) when it will be completed; and
e) how the results will be evaluated.

4 Support

4.1 Resources

This section of the Code relates to Clause 7 of BS 10008, ‘Support’.
The organization needs to determine and provide the resources needed for the establishment, implementation, maintenance and continual improvement of the information management system.

4.2 Competence

The organization needs to:

a) determine the necessary competence of person(s) doing work under its control that affects its information management performance;
b) ensure that these persons are competent on the basis of appropriate education, training or experience;
c) where applicable, take actions to acquire the necessary competence, and evaluate the effectiveness of the actions taken; and
d) retain appropriate documented information as evidence of competence.

NOTE: Applicable actions may include, for example: the provision of training to, the mentoring of, or the reassignment of current workers; or the hiring or contracting of competent persons.

4.3 Awareness

Workers doing work under the organization’s control should be aware of:

a) the information management policy;
b) their contribution to the effectiveness of the information management system, including the benefits of improved information management performance; and
c) the implications of not conforming to the information management system requirements.

4.4 Reporting and communications

It is important when developing policies and procedures to ensure that:

- information related to the policies and procedures is made available to those that are or may be affected by them;
- there is a mechanism for feedback from the implementers of the policies and procedures;
- there is a mechanism for reviewing risks related to the policies and procedures;
- details of any challenges to the authenticity and/or integrity of stored information is fed back to those responsible for compliance with BS 10008 and the Code;
- the methods used for these communications are regularly assessed for effectiveness, and updated where necessary.

KEY ISSUE

> Ensure that a reporting and communications mechanism is in place, to ensure that new or updated policies and procedures are implemented by all appropriate workers.

This documentation should be retained in compliance with the retention schedule.

KEY ISSUE

> Retain documentation related to reporting and communications procedures in line with the retention schedule.

4.5 Documented information

4.5.1 General

Documented information (also known as records) related to the process of managing information stored electronically needs to be created and retained for as long as necessary. Section 4.5.2 details procedural documentation that needs to be created and retained. This section also includes information related to the management of this information, including the requirement for version control and appropriate retention periods.

4.5.2 Procedural documentation

4.5.2.1 General

Compliance with BS 10008 and the Code requires the availability and use of specified documentation. This documentation consists of the:

- information management policy statement (see 2.2.2);
- information retention schedule (see 2.2.2.4.3.2);
- information security policy document (see 2.2.3);
- procedures manual (see 4.5.2.2);
- system description manual (see 4.5.2.3).

The availability of these documents, and demonstrable adherence to the procedures they describe, should, if effectively constructed, provide the audit trail that may be used to demonstrate the authenticity of information stored, and so enhance its evidential weight.

Note that each of the documents mentioned above may actually be maintained as multiple documents, or these documents may be combined. The key recommendation is that the documentation exists, is maintained and is readily accessible to those authorized within the organization to access it and to any authorized third party who may require access.

All documentation needs to be maintained in line with existing working practices, and so should be maintained under a version control system (see 5.7).

Additional documentation may be required to support the daily operation of the system, for example:

- the quality control log (see 5.2.3.7);
- the system maintenance log (see 5.15);
- an audit trail (see 4.5.3);
- compliance statements (see 6.3.4).

The content of the documentation described above can easily become unreliable where there are no procedures in place to ensure that they keep pace with both organizational and system changes. Unreliable documentation may adversely affect legal arguments relating to the correct operation of an information management system. It is therefore important to ensure that the definitive versions of system documents are brought under configuration management control, and are firmly linked to the organization’s change management procedures.

Where compliance with BS 10008 and the Code is claimed over a period of time during which different editions of the above documentation were appropriate, then all editions of this documentation should be kept, in compliance with the policy document. This is to ensure that, where information regarding the system at a point in the past is required, it can be obtained from this document store.

KEY ISSUE

> Documentation is essential evidence for the policies, procedures and technology used. It is part of the system audit trail.

4.5.2.2 Information management system procedures

4.5.2.2.1 General

The organization should maintain documentation (a procedures manual) for each information management system conforming to BS 10008 and the Code, describing all procedures related to the operation and use of the system, including input to, operation of and output from the system. This manual may include references to other controlled documentation as appropriate.

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.

This procedures manual should be readily accessible to all appropriate users of the system.

KEY ISSUE

> A procedures manual should be made available, containing details of (or reference to) other relevant documentation concerning all procedures relevant to the information management system.

4.5.2.2.2 Content

The procedures manual should include the topics listed below.

Topic	Content	Section of BIP 0008-1
Information capture	Importing or creating information to be stored	5.2
Self-modifying files	Managing files that can automatically modify themselves	5.3
Compound documents	Managing multi-file documents	5.4
Information in structured databases	Managing information stored in structured databases or other electronic systems	5.5
Big data considerations	Managing very large amounts of data in multiple computer systems	5.6
Version control	Managing multiple versions of a document	5.7

Storage systems	Technology used to store electronic information	5.8
Information retention	Implementing the retention schedule	5.9.1
Information disposal	Implementing the retention schedule and managing the information disposal procedures	5.9.2
Information transfer	Managing integrity during file transmission/transportation	5.10
Indexing	Creating index entries for captured information	5.11
Output	Creating authenticated output from the system	5.12
Identity	Managing the identity of individuals and systems	5.13
Information security procedures	Managing the security of information and systems	5.14
System maintenance	Maintaining physical equipment	5.15
External service provision	Using third parties for outsourcing	5.16
Information management testing	Ensuring systems meet business needs	5.17

Table 3 – Procedures manual topics

4.5.2.3 Key technology components

4.5.2.3.1 General

A description of hardware, software and network elements that comprise the system and how they interact, including details of system configuration, is needed. It 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, or by including a ‘change control’ section in the manual. What is important is that there is 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 introduction of BS 10008 and the Code cannot be considered as meeting its provisions unless the controls and procedures described in BS 10008 and the Code have been in place from the time of storing that information.

Where the information management policy requires compliance with particular national and/or international standards (see 2.2.2.4), the system description manual should include a section demonstrating compliance with those standards. This enables system auditors to check the performance and reliability of the system against these standards.

KEY ISSUES

> A system description manual should be made available, containing details of (or reference to other relevant documentation that contains details of) all technology-related issues relevant to the information management system at any point in time.

> Document any compliance to standards methodology implemented.

4.5.2.3.2 Content

The system description manual should include the topics listed below.

Topic	Content	Section of BIP 0008-1
Image processing	Use of image processing technology	5.2.4
Forms	Use of form overlays and electronic forms systems	5.2.4.6/5.2.4.7
Storage systems	Types of storage medium and format in use	5.8
Environmental considerations	Environment of technology installations	5.8.5
Media migration	Moving stored information between storage systems	5.8.7
Format conversion	Changing stored information from one file format to another	5.8.8
Compression techniques	Use of compression techniques	5.10.3.2
Business process management and workflow systems	Managing workflow systems	5.10.5
Access rights	Configuration of user access management tools	5.14.2
Encryption	Use of encryption technology	5.14.3
Digital signatures	Use of digital signatures	5.14.4
Other authentication techniques	Use of other authentication techniques	5.14.5
Audit trails	Details of available audit trail information	4.5.3
Date and time stamps	Accuracy of date and time	4.5.3.9

Table 4 – System description manual topics

4.5.2.4 Maintenance of documentation

Compliance with BS 10008 and the Code requires the availability and use of specified documentation. Maintenance of this documentation should be in accordance with the requirements of BS 10008 and the recommendations of the Code, including the keeping of records of this maintenance. This documentation should be subject to records management disciplines that are at least as good as those applied to the organization's other vital business records.

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

Documentation may be stored electronically in the information management system, subject to the same controls as included in BS 10008 and the Code, as paper or microform in secure locations, or as any combination of these.

KEY ISSUES

> There should be documented procedures for ensuring that documentation is up to date.

> Ensure that the version appropriate at a particular point in time in the past can be identified and accessed.

4.5.2.5 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 (see 4.5.2.2) is kept up to date, and that all previous versions are kept in compliance with the policy document (see 2.2.2).

KEY ISSUES

> All changes to operational procedures should be managed by a change control procedure, and include updating the procedures manual.

> Superseded versions of the procedures manual should be kept in compliance with the retention schedule, together with the dates during which they were in operation.

> The procedures manual should be regularly reviewed, to ensure that it is up to date.

> All changes should be reviewed to ensure that compliance with BS 10008 and the Code is not compromised.

4.5.3 Audit trails

4.5.3.1 General

When preparing information for use as evidence, it is often necessary to provide further supporting information. This information may include details such as date of storage of the information, details of movement of the information from medium to medium, and evidence of the controlled operation of the system. These details are known as 'audit trail' information.

This audit trail information is needed to enable the working of the system to be demonstrated, as well as the progress of information through the system, from receipt to final deletion. Audit trails need to be comprehensive and properly looked after, as without them the integrity and authenticity, and thus the evidential weight, of the information stored in the system could be called into question.

Audit trails may be but do not need to be automatically produced. They need to be designed so that an independent party (such as an auditor) is able to trace the history of an individual document, or the operation and management of the technology.

4.5.3.2 Purpose

The audit trail as defined for BS 10008 and the Code consists of the aggregate of the information necessary to provide a historical record of all significant events associated with the stored information and the information management system. As such it covers the answers to all the classic questions concerning the provenance of any piece of information stored within the information management system:

- Who?
- What?
- Where?
- When?
- Why?
- How?

These audit trail details can be split into two categories:

1. system (including the hardware platform(s), applications and operating software, configuration, and processes and procedures);
2. stored information.

In most organizations, the audit trail will consist of a collection of system- and operator-generated logs.

It is essential that system clocks be synchronized with an accurate time source[5] to ensure that times recorded in audit trails are consistent and reliable.

For further advice on audit trails and on building information management systems ‘fit for audit’, see PD 0018:2001, *Information management systems — Building systems fit for audit*.

4.5.3.3 Generation

Audit trail data should, as far as practicable, be generated automatically by the system, and the system description manual (see 4.5.2.3) should describe the processes. In this case, the data should be created and stored immediately following the event that is being audited.

Where audit trail data are not generated automatically by the system, procedures for their manual (or other) generation should be implemented. In this case, the data should be created as soon as possible after the event that is being documented. For example, if the record is of when an operator started work, the time should be recorded before work actually starts. If the record is of when preparation of a particular batch of documents was started, the time should be recorded just before the preparation of that batch commences.

It should not be possible to make any changes to the system or the information within the system without creation of an audit trail entry.

It should not be possible to amend audit trail data. Deletion should only be possible in accordance with the organization’s retention policy.

[5] See section on ’Trusted Time’ in BIP 0008-3 (2014).

KEY ISSUES

> Audit trail data should be generated automatically wherever possible, and should include entries for all system and data changes.

> It should not be possible to alter audit trails.

4.5.3.4 Content

4.5.3.4.1 General

The audit trail content is critical, as it can be used to determine the relevance and importance of particular document types. Audit trails may also be needed to audit such activities as creation, access and deletion. Thus, the audit trail needs to include a record of all relevant occurrences. If any significant occurrence is not audited, then the whole audit trail can be discredited and, as a direct result, all or any information held within the system will also be able to be discredited.

Thus, technologies for electronic information storage should be chosen with audit trail requirements in mind. This may result in technologies being rejected that may otherwise have appeared suitable for a particular application.

KEY ISSUE

> When choosing electronic information storage systems, consider suitable audit trail functionality as a basic system requirement.

Audit trails should include details related to the following procedures:

- information capture (see 5.2);
- batch processing (see 5.2.3.9);
- indexing (see 5.11);
- version control (see 5.7);
- disposal of information (see 5.9.2).

All audit trails should include details of who the individual was or what application/system component was responsible for the audited event.

4.5.3.4.2 Information capture

Audit trail data about the information capture process provides invaluable information to assist in the authentication of stored information. Details such as capture time, operator and capture device may prove vital when authenticity is challenged.

Information that may be stored in the audit trail will typically include:

- document or file identification;
- process date and time stamp;
- batch reference (for batch input);
- number of pages (for document scanning) or data records (data capture);
- quality control check approval;
- an identifier for each document or file that was indexed;
- worker and workstation identifier;
- final write to storage.

The choice of actual data to be stored in the audit trail will depend upon the application and the system.

KEY ISSUE

> Records should be kept in the audit trail of key information concerning information captured by or imported into the system.

Information may be captured by the system on a file-by-file basis. This is particularly relevant where data files are imported into the system. In this case, the following audit trail information may be stored:

- a unique file identifier;
- number of documents/pages/data records within the file;
- size of the file (e.g. kilobytes);
- file format.

Where compound files are captured, audit trail data should include the relationship between the parts.

KEY ISSUE

> Where individual files are captured, audit trail data should be kept on a file-by-file basis.

4.5.3.4.3 Batch processing

Where batch processes are involved, it is common to use batch control documents to record and reconcile batch content and process completion. These form an integral part of the overall audit trail.

Where data are captured on a batch basis, particularly in document scanning applications, the following audit trail information should be stored:

- unique batch identifier;
- operator identifier;
- type of material scanned (e.g. paper documents, roll microfilm and/or aperture cards);
- quantity of material in the batch (e.g. number of documents, number of pages (single/double-sided) or number of microfilm frames);
- details of image processing performed during the scanning process, where this is different from any default image processing described in the system description manual.

Audit trail data should be stored so that it is easy to confirm that:

- all required activity has been performed for that batch;
- details of any anomalies or discrepancies that have been encountered (e.g. number of pages written to storage not agreeing with number of pages scanned) have been recorded and dealt with;
- quality control procedures have been completed;
- required exception processing has been completed.

KEY ISSUE

> When working in batches, audit trail data should include relevant information about batch reconciliation.

4.5.3.4.4 Indexing

Indexing information is vital to the information retrieval process, and so its accuracy is key to establishing the authenticity of stored information. Audit trail information detailing the creation and modification of indexes can be used to demonstrate that indexing procedures have been correctly followed.

It may be appropriate, depending upon the application, to keep ‘before’ and ‘after’ records of index entries that are amended or deleted.

Where an index item relates to deleted information, this fact should be documented.

KEY ISSUES

> A record should be kept in the audit trail of the date and time of the creation, amendment and deletion of every index file.

> Audit trail data should include an identifier for each document or data file that is indexed.

4.5.3.4.5 Version control

Where a change is made to a data file, audit trail data should be created and stored, identifying the nature and details of the change. Such changes could have been made by an individual, or automatically by the system.

Where appropriate, previous versions of data files should be referenced in the audit trail data, to identify the nature of the change.

KEY ISSUES

> Where changes are made to stored information, audit trail data should be stored detailing the change.

> It may be appropriate to keep a copy of the information before the change, to document the process that has been carried out.

4.5.3.4.6 Disposal of information

Records should be kept in an audit trail of the disposal of information at the end of the relevant retention period. This is particularly significant if it is necessary to demonstrate that information disposal was in accordance with the retention and disposal policy and was not undertaken to hide information that could have been incriminating evidence.

Where original documents are destroyed following a scanning process, records should be kept in an audit trail of the destruction of the original documents.

KEY ISSUE

> Keep records of any information disposal procedures carried out.

4.5.3.5 Security of audit trails

If the authenticity of stored information is questioned, the integrity of the audit trail may be fundamental in establishing the authenticity, and thus the evidential weight, of the stored information. If the possibility exists that the audit trail data could be modified, this will reduce the evidential weight of any information to which these records apply.

The audit trail should be kept at the level of security appropriate to preventing any change to any data within it, and in accordance with the organization’s information security policy (as well as the retention policy).

The audit trail should be subject to internal records management policies and procedures that are at least as good as other ‘vital records’ of the organization.

Secure backup copies of the audit trail should be kept, including automated and manual audit trail data.

Where file recovery procedures have been implemented, sufficient audit trail data should be stored to demonstrate that the recovery did not affect information authenticity.

For least risk, store audit trail data on WORM media. If a rewritable medium is used, then additional procedures need to be implemented to reduce the risk of changes being made. The use of magnetic tape will make it relatively difficult to modify data, as magnetic tape is normally a serially written medium.

If audit trail data has been modified, then any such modification should be audited.

Paper documents used for audit trail data should be frequently removed from the place of use and stored securely. The longer a document used for audit trail data (e.g. operator logs) is left in a relatively insecure place, for example, at a workstation, the higher the risk of tampering. Users need to assess such risk when using paper for audit trail records. Where paper documents are used, storing copies of them on the information management system is recommended.

KEY ISSUES

> Wherever possible, store audit trail data in an unmodifiable form.

> Where this is not possible, use security measures to ensure that it is not modified.

4.5.3.6 Management

Audit trail information needs to be properly managed, as it may be of critical importance to the organization. All claims of compliance with organizational policies may be discredited if the audit trail is not treated correctly and cannot be interpreted unambiguously.

Where audit trail data are stored manually by an operator, it may be impractical and unnecessary to create audit trail data on a per document basis. For example, when undertaking document preparation for scanning, it may be sufficient to document the date and, where appropriate, time at which preparation of a batch of documents started and ended. It may suffice to document simply when the operator started and ended work, provided it is possible to identify subsequently which operator prepared which documents.

KEY ISSUES

> Ensure that the audit trail data are authentic, accessible, sufficiently comprehensive and understandable.

> Audit trail data should include a date and (where appropriate) time stamp.

> It may be appropriate to audit at the batch level rather than the individual document level, in some parts of the capture and storage processes.

4.5.3.7 Storage and retention

The storage of audit trail data is a topic often not included in an organization’s information management policies. As they are frequently created automatically, and infrequently accessed, they are usually forgotten, and thus not subject to adequate control.

To ensure that all relevant audit trail data are stored, 'audit trail data' should be included as a specific document type in the policy document. It should be stored for at least as long as the information to which it refers.

Some systems control the size of audit trail data files by the use of 'looping', which sets the maximum size for the data file, and when this size is reached new data overwrite the oldest data in the file. Thus, old audit trail data are lost. This process may not be in conformance to required retention periods.

There should be procedures that identify circumstances when an audit trail data file becomes full, and action is taken to retain data as required by the retention policy.

Where an organization is working within a BS EN ISO 9001 environment, typically audit trail data relating to compliance with the quality management system are destroyed after a short period of time. This is not the case with audit trail data from information management systems, which should be stored for the same period as that of the data to which it relates.

KEY ISSUE

> Ensure that audit trail data are retained for at least the same period as that of the data to which they relate.

4.5.3.8 Format

The format of audit trail data is a topic often not included in an organization's information management policies.

Frequently, when an organization wants to automate its computer operations environment, it makes use of operating system logs to monitor the system for specific events or error conditions.

At an application level, ensuring that the application provider uses standard error messages, typically agreed with the organization during the design stage, also enables application status conditions to be monitored.

For example, if an application reads invalid data from a file, rather than just aborting the program with nobody aware of what has happened (until the users raise a support call), if the program writes a status message to an error/system log, in an agreed format, the monitoring software will detect this and notify the user and/or support staff.

These notifications are an important trigger to investigate the continued, proper operation of the system. The entries into the error/system log that caused the monitoring software to raise the alert should be part of the audit trail and should be controlled in accordance with BS 10008 and the Code.

KEY ISSUE

> Use audit trail formats that enable easy interpretation, both by system users and by automated monitoring tools.

4.5.3.9 Date and time stamps

Being able to determine the date and/or time of an event (such as the capture of a document by the system) can be an important piece of evidence. Thus, all appropriate events should be date- and/or time-stamped. Depending on its importance, date and time information may be stored on a batch or individual event basis.

Where accuracy of date and/or time clocks is important, regular checking of system clocks should be carried out. Any errors should be corrected and any actions taken documented. Only authorized personnel should be able to change system clocks.

Where clocks are changed on a seasonal basis, for example 'summer time', then procedures to be followed should be documented. Clock changes may be automated, but it is important to check that the appropriate action has been completed at the right time.

In some applications, the clock may be confused by multiple time zones. Where this may be an issue, relevant procedures should be implemented to ensure that the appropriate time information is accessed.

Where there is a particular need to demonstrate the accuracy of date and time stamps, the use of trusted third-party services for this may be considered. Such services can also be used to demonstrate that system clocks have not been tampered with. 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.

KEY ISSUES

> Mark all appropriate events with date and/or time stamps. Check system clock accuracy from time to time.

> Where multiple time zones are in use, relevant controls need to be implemented to manage time stamps/clocks in an appropriate manner.

> Assess the value of Trusted Time, and use where appropriate.

4.5.3.10 Access and interpretation

Access to the audit trail information needs to be controlled. In some applications, access may only be needed infrequently, and thus it is important that the interpretation procedures are documented. As audit trail data may be inspected by authorized external personnel (such as auditors) who have little or no familiarity with the system, interpretation procedures should be understandable by non-technical users.

There are frequently a number of departments (or individuals) within an organization (or external to the organization), including those representing user, audit and legal functions, that may need access to specific parts of audit trail information. This access should be controlled, to reduce the possibility of compromise.

Access to the audit trail should itself be recorded in the audit trail.

KEY ISSUES

> Relevant personnel should have access to audit trail data commensurate to their role.

> There should be documented procedures that are followed when audit trail data needs to be accessed and interpreted.

5 Operation

5.1 Management overview

This section of the Code relates to Clause 8 of BS 10008, 'Operation'.

When responding to questions about the authenticity of stored information, one of the major issues that will need to be resolved is 'Was the system operated correctly at all relevant times?' In order to be able to deal confidently with this issue, all relevant procedural and technology issues will need to be well thought out, be complete in their scope, and be operated by competent individuals. It is also essential that, because of the long time intervals involved, all relevant procedures are documented. This documentation is described in BS 10008 and the Code as the procedures manual and the system description manual, which document all the operating procedures and technology that the organization needs, where appropriate, to implement.

5.2 Information capture

5.2.1 Information loss

Whenever information is 'copied', 'converted', 'moved', 'updated' or 'captured' within a process, there is the potential for loss of some of the information. For example, when a paper document is scanned or photocopied, resolution may be lost, resulting in the illegibility of small print. Or when an electronic document is converted from one format to another, some metadata may be lost.

In some applications, where original paper documents are scanned as part of an electronic storage system, there may be a process to destroy the original document. This process will also result in information loss. In this case, the information may be the type of paper used, the ink in the pen used to sign the document or a fingerprint on the page.

In all these cases, the organization should review the information loss (potential or actual) resulting from each type of transformation process, and make a decision as to whether the loss is 'acceptable', and that no 'material' information is lost. This decision will need to be based on a risk assessment, reviewing the impact of the potential loss of evidence against the cost of retaining that evidence.

KEY ISSUE

> Review any information loss by each transformation process, and make a business decision on the acceptability of the resultant documents, based on a risk assessment.

5.2.2 Creation and importing

Documents containing information may be created by the information management system, or imported into it. The authenticity of documents at the time they are imported is of critical importance when they are later called upon and their authenticity and integrity are subject to scrutiny.

Documents can be stored in two forms, in either image or data format. In either form, they can be imported into the information management system in a variety of formats.

Image formats are typically obtained from:

- paper documents (originals, photocopies and faxes);
- automatic facsimile entry (via a fax server);
- microfilm and microfiche.

Image formats are typically bitmaps of an original document. Image formats can also be obtained from documents in data format. Details of procedures for capturing documents in image format are discussed in 5.2.3.

Data formats store information in the original format, typically requiring the original software to retrieve the information. There are a number of ‘standard’ formats that can be retrieved by many software packages (e.g. text files or comma separated delimited files). Examples of data formats are:

- office systems such as word processors and spreadsheets;
- ’standard’ file formats such as PDF;
- CAD drawings;
- email messages;
- electronic data interchange (EDI) files;
- HTML and web pages (intranet and internet);
- instant and SMS messages;
- Extensible Markup Language (XML) messages;
- databases and other data stores;
- program to program data files.

In all cases, the information contained in the data can be directly accessed by the read software. Details of procedures for capturing documents in data format are discussed in 5.2.5.

NOTE: It is also possible to have electronic documents in mixed image and data formats (e.g. a letter in a text format with an embedded bitmapped signature).

Where an image or data file originates from outside the boundaries of control of the organization employing the information management system, there may be little or no control over, or knowledge of, the procedures or processes involved in the production or authorization of the image or data file. In these circumstances, the organization will need to take care that the document is what it purports to be, that it has not been tampered with and that the identity of the originator is genuine. The level of checking of these criteria will depend upon the nature of the particular document in question.

Such boundary situations can also exist within an organization. In these circumstances, the part of the organization with the information management system should not assume that an image or a data file is what it purports to be, simply because it came from another part of the same organization.

KEY ISSUES

> There should be documented procedures for the creation or importation of image or data files into the system.

> There should be documented procedures for the confirmation of the integrity and authenticity of image or data files that are created by or imported into the system, at the time of creation or importation.

5.2.3 Document image capture

5.2.3.1 Introduction

Document images may be obtained from original (or copy) documents stored on many different media types. Paper or microform documents are common, but images can be imported into the information management system in a wide variety of ways (e.g. facsimile, microform, email, email attachment or file transfer).

If the information management system is used for storing document images, the procedures involved in the capture of those images should be documented. This section contains recommendations relating to the procedures relevant to document image capture using scanning systems.

Where paper document input is concerned, these procedures may include:

- document preparation;
- document batching;
- photocopying;
- scanning;
- image quality control.

5.2.3.2 Preparation of paper documents

All paper documents need to be examined prior to the scanning process, to ensure that as high a quality image as possible is obtained. Attributes such as their physical state (thin paper, creased or stapled, etc.), attributes of the information (black and white, colour or tonal range, etc.), paper size, weight and binding, and print colour can all affect the physical scanning process.

Where documents are found that 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 (see 5.2.3.10) or transparent wallets could be used.

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 an original document has physical attachments, for example, stick-on notes, a procedure for handling these, making sure that all relevant information is captured, should be implemented. This might be achieved, for example, by capturing a separate image of the attachment, with appropriate index data to associate it with the source page. If only a single image is captured with the attachment in place, the index 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 original document, it might be preferable to ensure that an image of the original document without the attachment is captured.

Where an original document has physical amendments, for example where correction fluid has been used, the information management system should ensure that the presence of such amendments is noted.

Where it is evident that the document being prepared for scanning is a photocopy, and this fact may be important in the future, it should be clear to a user on any subsequent retrieval of the scanned image that this was the case. This is frequently the case when procedure would normally require receipt of or access to the original document (e.g. a birth certificate), but in a particular case a photocopy was accepted.

There should be procedures in place to ensure that all pages of a multi-page document are kept together and in the appropriate order before, during and after scanning.

KEY ISSUES

> There should be documented procedures for the examination of paper documents prior to scanning.

> There should be documented procedures for the handling of paper documents that may cause scanning difficulties.

> There should be documented procedures for the removal of staples and other binding methods.

> There should be documented procedures for dealing with stick-on notes or other attachments.

> There should be documented procedures for dealing with paper documents that have been physically amended.

> There should be documented procedures for marking documents that are evidently photocopies.

> There should be documented procedures for ensuring that the integrity of multi-page documents is maintained.

5.2.3.3 Paper document formats

5.2.3.3.1 Line drawings/art

For line drawings/art that form part of otherwise text-oriented documents, the scanning resolutions applicable to text are typically also satisfactory for the drawings. With printed material, where fine lines are used in the artwork, 200 dots per inch (dpi) may be too low, but this can only be determined via tests on sample documents.

5.2.3.3.2 Handwritten material

With material where a modern ball-point pen or pencil was used, often 200 or 300 dpi will be adequate, but lower resolutions may be acceptable. But for older material where a steel-nibbed pen was used (e.g. to produce ‘copperplate’ handwriting), the thinness of the upstrokes may often be such that higher resolutions (such as 400 dpi) will be the minimum that will satisfactorily capture the text without fine detail in the upstrokes being lost.

Handwriting (or hand drawing) using pencils can be faint, and difficult to reproduce. Care should be taken when scanning to ensure that image brightness and contrast are appropriate for these images.

5.2.3.3.3 Plans and drawings

For hand-drawn architectural and engineering drawings, there may be finer lines present than would be the case with a typical ‘full-sized’ CAD drawing and, although 200 dpi will usually be a satisfactory resolution, tests should be done to ensure that the finest detail is captured. It may prove necessary to use 300 or even 400 dpi.

With plans and drawings, line thickness will vary depending on the size of the output. With the larger formats, A0 or A1 (or US or imperial equivalents), line thickness tends not to be as fine as with some hand-drawn material. Tests should be undertaken to determine the appropriate parameters for scanning, with the appropriate dpi chosen to ensure a satisfactory reproduction of the drawing. Often it will be satisfactory to scan at 200 dpi; however, with smaller plans and drawings, higher resolutions may be required.

If the scanning is to be done from low quality copies of the originals, and if these copies have been reduced from the originals (which is quite common), then a higher resolution may be required than would otherwise have been satisfactory.

With drawings, dimensional accuracy may be important. Because drawings are often large, the paper or film may undergo dimensional change (due mainly to variations in moisture content). For working drawings it is often a requirement when scanning that dimensional inaccuracies are corrected, that is the scanned image may be post-processed to correct scale inaccuracies, skew or lack of orthogonality. Such corrections mean that the subsequent image is not a true facsimile of the original. In cases where the legal admissibility of such images might be called into question, it would be prudent to preserve an uncorrected version of the scanned image as well as the corrected version.

5.2.3.3.4 Maps

With maps, a minimum resolution of 400 dpi will often be required, but much higher resolutions (e.g. up to 1,000 dpi) may be required with some material. Tests should be performed to identify the appropriate resolution that will produce the appropriate quality results.

As with drawings, scanned images of maps are frequently corrected for scale inaccuracies and lack of orthogonality after scanning.

Where coloured maps are being scanned, and the colour is to be preserved, the scanner should be capable of capturing individual colours with the required discrimination. While the number of colours subjectively present may be quite small, 8-bit colour (256 colours) may be inadequate and it may be necessary to scan with 24-bit colour in order to provide the required colour discrimination. Tests should be done to determine how many ‘bits’ of colour are required.

5.2.3.3.5 Half-tone material

Where half-tone material (black and white or colour separated) is present on a page along with text and/or line art, the objectives of the scanning should be addressed.

If the objective is to produce a scanned image that is comparable in quality to a ‘normal’ black-and-white photocopy, then a scanner that produces a digital image in black and white will suffice. The resolution may have to be higher than that which would be acceptable for text only: where 200 dpi produces a poor quality image of the half-tone material; 300 dpi will give a better quality image; whilst in some applications 400 dpi may be appropriate.

Most scanners have different settings for text or line art originals and for originals with half-tones. The scanner settings that are optimal for text or line art originals are far from optimal for those with half-tones, and vice versa. Where it is important to reproduce half-tones, the half-tone scanner settings should be used. Where half-tone content has ‘cosmetic’ value only and does not contribute to the essential information content of the original, then the text or line art scanner settings should be used.

If the half-tone material is to be captured to a quality level comparable to that of a typical (good quality) photocopy, then there are two options. One option is to scan the document with the scanner settings set at ‘normal’, at a higher resolution than would be necessary for the text alone. The other option is to scan the document twice, to create two images, one where the text/line art is captured to satisfactory quality and the other where the half-tone material is satisfactorily captured. In the latter case a record should be kept that the production of the two images involved different scanner settings (affecting the processing performed on the images).

If the half-tone material is to be reproduced to a quality comparable to that of the original, then it should be processed according to the recommendations for photographs.

5.2.3.3.6 Continuous-tone images

Continuous-tone images include photographs, medical and industrial radiographs (X-rays) and images generated by computer as photographic style images, including ultrasound, computed tomography (CT) and magnetic resonance (MR) images.

With material containing continuous-tone areas (grey scale or colour), where the tonal information should be preserved, scanning should be performed with a scanner capable of capturing the required number of grey levels and/or colour. The number of levels that is appropriate should be determined by benchmark tests on the sample set of documents.

For images from photographic material, the number of grey levels will typically be 16, 64 or 256 (i.e. 4, 6 or 8 bits per pixel). For very high quality images, 256 levels are normally used, and for X-rays, up to 1,024 levels of grey (i.e. 10 bits per pixel) may be necessary.

For colour photographs, 24 bits per pixel of colour information is used in most applications, but for very high quality images, up to 36 bits per pixel may be necessary. Typically, 15 or 16 bits of colour are used; for source material containing only a small palette of colours, 256 levels may suffice. Tests should be performed to determine how many colour levels are required.

With continuous-tone colour, most scanners capture 8 bits of colour information in three different regions of the colour spectrum: red, green and blue (‘RGB’), resulting in 24 bits per pixel, or the ability to reproduce over 16 million colour variations. With only 8 bits of colour information (256 levels), however, there may be a noticeable ‘blockiness’ in the image if the original contains a broad range of colours.

Scanning resolution requirements for documents containing colour are normally similar to those for black-and-white material, particularly if there is text present on the original. Thus scanning may be performed at 100 to 400 dpi. If there is no text present on the original, satisfactory images may be achieved at lower resolutions, down to television quality levels (about 350 lines per image frame); this would typically be satisfactory for identity photographs and similar applications.

To assess image quality, in general it is satisfactory to compare the screen image with the original. If there is likely to be use of high quality hard copy images then the comparison should be made between hard copies of the images, produced on a high quality colour printer, and the originals.

Care should be taken when comparing screen colours with an original that the colours were correctly balanced at the time of image capture, and that the display system has also been calibrated correctly. Otherwise the displayed colours may be different from the colours on the original. Care should also be taken that the colours were correctly balanced at the time of image capture, when comparing the original with hard copies of the captured image.

Where colour accuracy is important, a standard colour gamut test chart should be scanned at the same time as the original (or batch of originals scanned at the same time), and the image of this chart stored along with the original.[6]

5.2.3.3.7 Mixed mode documents

Mixed mode documents comprise more than one document type inside a single document (e.g. photograph and text). From a scanning perspective, documents containing half-tone material are essentially of this type, even though the original has been created in a single print operation. As described in 5.2.3.3.5, the use of scanner settings optimized for one type of material can result in the loss of information in material of other types. As previously suggested, one solution is to capture multiple images, with scanner settings (or even scanner type) selected to optimize the image quality for each material type.

One option is to use a scanning system that can scan mixed mode documents automatically, with automatic detection of each type of material and automatic optimization of the settings for each type. These systems can also be set to select the most appropriate compression algorithm for each type of material. Benchmark testing should be done to ensure that the results are acceptable.

[6] Rochester Institute of Technology, Process Ink Gamut Chart. Available from Rochester Institute of Technology, T & E Center, One Lomb Memorial Drive, Rochester, New York 14623, USA

5.2.3.3.8 Documents with note sheets attached

Some documents may have note sheets or notelets attached. Care should be taken when scanning such documents. It may be necessary to remove the attachment where, for example, it obscures information on the document. If removal is required, the note should be marked or stamped as being a part or page of the document to which it was attached, and scanned and indexed separately. The original page should also be indexed to indicate that it has an attachment.

Where a system has a facility to indicate that a document has a related image, this facility should be used.

5.2.3.3.9 Microform documents

Microforms should be examined carefully prior to deciding upon the scanning approach.

Within multi-frame microfilm media (roll film, microfiche, microfiche jackets and multi-frame aperture cards), unless the inter-frame gap can be detected unambiguously, automated frame detection should not be used.

If the gap is not detected, multiple frames may merge into one image. Depending on the physical characteristics of the scanning system it is possible that some part(s) of the digitized image may be lost.

With jacketed film, film strips may overlap. The processing procedures should ensure that such overlaps are detected and corrected before scanning, otherwise some page images will be missing or illegible, in whole or in part.

Where a rotary camera has been used, images on the film may not have a one-to-one correspondence with the original documents. For example, two pages may have been fed at once, so that on the film, part or all of an original page may be missing.

5.2.3.4 Scanning processes

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

Where scanning procedures may be varied for particular documents, there should be methods for ensuring that the appropriate scanner settings are used for each document. Examples of document attributes that may require different scanner settings are:

- double-sided vs. single-sided;
- colour scanning vs. black-and-white scanning;
- various text and/or background colours.

In many applications, it is important that all original documents in a particular batch are scanned. This may be achieved by comparing a count of images created with the number of documents in the batch. Alternatively, each original document could be marked during the scanning process. These marks need to be checked to ensure they are present on every document. Where batching is not used, alternative procedures for ensuring that all documents are scanned may be needed.

When using the count of images created, take into consideration any blank pages that may be removed by automated processes (see 5.2.4). It is preferable to use a count of images before the blank pages are removed.

Many scanners have automatic document feeders (ADF) that can reliably detect document mis-feeds, so minimizing the risk that a document may pass through the scanner without being scanned. If such devices are not used, procedures are required to ensure that the scanner operator manually handles every document in order to reduce the probability of specific documents not being scanned.

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

EXAMPLE

An organization has decided to scan incoming post. One of the characteristics required of the scanning process is to be able to demonstrate that all post has been scanned.

Hence, a process of opening each envelope carefully, batching the documents with the relevant separators inserted, and counting the pages to be scanned was used. This count was then compared with the image count from the scanners. In order to reduce the risk of manual counting errors, and to ease the identification of specific errors, a small batch size was used (about 50 sheets).

In practice, it was determined that the manual count was more likely to be incorrect when a mismatch occurred. However, the importance of the requirement to be able to demonstrate that all documents had been scanned meant that the process of counting pages was continued.

Where large original documents are 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 section of the Code on expungement (see 5.9.2). This unique identity could be a system-generated sequence number, which may be used for internal control purposes only.

EXAMPLE

Double-sided application forms were being scanned by a financial institution, on a double-sided scanner. To avoid capturing ‘blank’ pages, a process was put into place to check the reverse side of every form manually prior to scanning and to separate those out without anything entered. The forms were then scanned either single- or double-sided as appropriate.

On investigating this process, it was discovered that the occasional reverse side of a form with a small amount of added information was being put into the single-sided selection. Thus, this information was lost on the electronic copy. It was also noted that a lot of resource was being used in the splitting exercise. It was decided to scan all forms double-sided, to ensure that all appropriate evidence was captured. The electronic storage overhead was considered to be acceptable in this case.

KEY ISSUES

> There should be documented procedures for scanning, including where scanner settings are varied for different document attributes.

> Checks should be made to ensure that every document is scanned, including when scanning double-sided documents with simplex scanners.

> There should be documented procedures for scanning large documents, so that no information is lost.

> There should be documented procedures that allocate a unique unalterable reference (e.g. a sequence number) to each scanned image.

5.2.3.5 Scanner resolution

Scanner resolution is typically measured in dpi. The higher the resolution, the finer the detail captured (i.e. the smaller the character that can be captured). However, the higher the resolution, the larger the image file size will be. When choosing the scanner resolution, a balance will need to be achieved between levels of detail and file size.

Where automated data capture is used (see 5.2.6), higher resolution images will typically lead to higher accuracy levels. So, when deciding upon the resolution that is required, one of the factors to take into account is the accuracy levels obtainable.

The resolution required for accurate optical character recognition (OCR) may need to be higher for colour or grey scale images than that for black-and-white text only documents. Also, the required resolution will depend upon the font style and font size used in printed text. Where lossy compression is used (see 5.10.3.2) the compression ratio used will need to be taken into account.

> EXAMPLE
>
> An organization decided to scan all documents at 300 dpi, in order to obtain high accuracy levels with its OCR technology. Before storage to the document repository, the images were ‘degraded’ to 200 dpi, to reduce file sizes. Tests of the degradation process (which were kept) demonstrated that no loss of human readability occurred.

KEY ISSUES

> Choose the scanner resolution (dpi) to be used with care, balancing fine detail capture with file size. Run tests on typical documents to identify the appropriate resolution setting.

> Where techniques such as OCR are used, increasing the scanner resolution will usually improve recognition rates. Using grey scale or colour scanning may also improve results.

5.2.3.6 Text

Typically, a scanner resolution of less than 100 dpi will not produce an image of acceptable quality. Detail may be missing from some characters, particularly if they contain thin elements, including serifs. Fonts under 6 point on the original may not be adequately reproduced.

With material containing particularly small type sizes (e.g. superscripts and subscripts), a resolution of 200 dpi, 300 dpi or higher may be necessary.

For material that may be processed using optical (or ‘intelligent’) character recognition (OCR/ICR), it may be beneficial to scan at a higher resolution than would be satisfactory for visual legibility. For example, while for much material, 200 dpi would be satisfactory for visual representation, it may be preferable to use 300 dpi resolution if OCR/ICR is to be used; similarly, where 300 dpi may be visually satisfactory, 400 dpi may be better for OCR.

No decisions should be made regarding choice or resolution without conducting tests.

Depending on the scanner, the image quality achieved from scanning at 100 dpi may not be greatly inferior to the quality at 200 dpi or 300 dpi.

It is important to bear in mind that a typical screen used for viewing document images has an effective resolution of about 100 dpi, or even less. This is normally adequate for much typed material, but ‘zooming’ may be required when viewing images with small print, and this in turn requires that the scanning resolution should be substantially greater than the basic display resolution.

While post-scan image enhancement may improve the subjective image quality and legibility, this should only be performed after careful tests to ensure that the resulting image remains an effectively 'true' facsimile of the original.

These tests should use a sample set of documents, and hard copies should be made of scanned images with and without image enhancement being used.

There should be no anomalies introduced into the enhanced image that are visible under normal office lighting conditions.

As it may be necessary to refer back to the test results in the future, the results should be stored in an appropriate manner.

5.2.3.7 Quality control

5.2.3.7.1 General

Procedures are required that reduce the risk of scanned images being of unsatisfactory quality. The evidential weight of scanned images will be increased if it can be demonstrated that the images are of good quality, and that the scanner was working to agreed standards at the time of scanning. This can be achieved by:

- ensuring that the scanner is working properly, by using test methods;
- checking the quality of every scanned image.

5.2.3.7.2 Scanner testing

Document scanners can be tested by using specific original documents, and examining various attributes on the scanned image. These original documents can be either 'typical' documents identified by the organization, or can be specific test targets designed for the purpose.

Quality control criteria may cover:

- overall legibility;
- smallest detail legibility (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, and 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 should be documented and agreed by all parties whose use of images is likely to be affected by image quality, including internal and external users. The agreed criteria should be realistic given the nature of the original documents and the characteristics of the scanning equipment.

Where the quality control procedures involve sampling of the scanned images, 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 original documents. The proportion sampled should be identified based on a risk assessment. For example, when starting scanning initially a relatively large sample may be selected (e.g. 10 per cent), which may be reduced (e.g. to 5 per cent or even lower) once the consistency of meeting the required quality standards has been demonstrated.

Scanner quality control checks should be used regularly to check that output image quality is acceptable. The frequency of scanner quality control 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 scanning system. Initially, it may be appropriate to scan a test target after every few thousand pages scanned.

5.2.3.7.3 Evaluating image quality

Procedures for the evaluation of image quality on a day-to-day basis should be documented. These procedures should include details of the evaluation and documentation of results. Where appropriate, results of quality control checks could be stored in a quality control log.

Care should be taken when assessing the quality of an image, as the viewing method (screen or print) can significantly affect the results obtained. When designing quality assessment procedures, standard viewing conditions should be used for quality assessment.

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

Where dimensional accuracy is important, procedures should be documented for checking that dimensional information is reproduced within acceptable tolerances. 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.

Where batching is used in the scanning process, there are advantages in relating quality control procedures to the batching process, enabling acceptance or rejection of one batch independently of any other batch.

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

5.2.3.7.4 Sample set

A sample set of original documents, or of documents equivalent in characteristics to the original documents, should be assembled for the purposes of evaluating scanner results against agreed quality control criteria. The chosen documents should be representative of the operational documents to be scanned. They should include examples of original documents whose quality is poor relative to those of the majority of the documents.

Image quality criteria when using a sample set of documents will mainly rely upon a subjective assessment of readability. Test documents with a range of character sizes typical of operational documents should be used wherever possible.

5.2.3.7.5 Scanner test targets

Scanner test targets have an advantage over sample original documents when checking quality, as they are designed to be objective rather than subjective. It is thus possible to agree specific attributes of a test image that will result in good quality scanned images from operational documents.

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

Regular use of test targets can show whether the scanner is performing consistently and in accordance with its specification. The frequency of testing will depend upon the volume of documents being scanned, the type and physical state of the original documents, and the type of scanner being used. For high volume applications, daily or more frequent testing should be used. Testing frequency should be increased if performance and/or quality issues are found. By storing the test target images, taken at specific dates and times, on the information management system, quality criteria can be reassessed at any time in the future, should the need arise.

An example of a scanner test target is given in BS ISO 12653 series, *Electronic imaging — Test target for the black-and-white scanning of office documents*. BS ISO 12653-1 specifies how a target is to be manufactured. BS ISO 12653-2 details the scanner attributes analysed by the target. BS ISO 12653-3 details a test target for use with scanning systems used at 300 dpi or lower.

These include:

- legibility;
- resolution;
- thin line detection;
- coverage of an A4 page;
- grey scale reproduction;
- dimensional accuracy.

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.

KEY ISSUES

> Quality control procedures should be used to ensure good quality scanned images.

> Agree quality control criteria with all relevant people. Use sample documents or scanner test targets.

> Record results of quality tests (or store images of test targets).

> Take care with viewing conditions of test images, as these can significantly affect the results obtained.

5.2.3.8 Rescanning

Some scanned images will need to be replaced, typically due to poor quality. Procedures for rescanning documents should be developed that ensure that images resulting from rescanning replace the original scanned image, and that batch numbering and audit trail procedures are not compromised.

KEY ISSUE

> Develop rescanning procedures to ensure that the audit trail is not compromised.

5.2.3.9 Document batching

There may be advantages in managing the scanning of paper documents using a batching process. This may make it easier to control individual documents, and to be able to perform quality control and other procedures on a sampling basis. Document batching can also be used for checking procedures, for example, to check that the number of images obtained from a batch equals the number of original documents scanned (see 5.2.3.4).

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, say 100 pages, then the documents in a single file cover may constitute a batch. If the file covers contain relatively few documents, say on average 10 pages, then a batch may consist of documents from more than one file cover. Choose the batch size so that it is not bigger than can be easily managed, nor smaller so that checking quality by sampling on a batch basis would result in significant process inefficiencies.

For some applications, however, the size of a batch may not be easily definable. In these cases, a batch may be defined in terms of how many documents have been scanned during a specified time period. Thus, for example, one batch could consist of all documents input during an hour or a day.

For some applications (especially where a workflow is implemented), where batching cannot be applied, alternative methods for controlling the scanning processes should be established. In these cases, every document may have to be controlled separately, to ensure that it is scanned to acceptable quality standards.

EXAMPLE

When scanning mail received, to ensure that all received documents were scanned, a document batching process was introduced, with a 'batch' size of approximately 50 sheets of paper. A separator sheet (containing a 'new document' bar code) was inserted at the beginning of the contents of each document (where in this case a document was all the contents of an envelope). The number of sheets of paper in each batch was counted, and written on a batch header sheet, together with operator name and date/time information. This value was used later to check that the correct number of images was obtained during scanning.

KEY ISSUE

> Paper documents should be grouped into batches of a known batch size (size to be assessed by experimentation).

5.2.3.10 Photocopying

It may be helpful for some documents to be photocopied, and the photocopy to be scanned. Where this has been done, users of the scanned image should be made aware that the image was obtained from a photocopy of an original document. This is to ensure that an image may be correctly identified as a true facsimile of an original 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. Stamping or marking the document as a 'photocopy' before it is scanned may achieve this, for example.

Examples of documents that might benefit from being photocopied before scanning include:

- documents that may be adversely affected by the scanning process, such as damaged or delicate documents (a flat-bed scanner rather than an ADF may help in this situation);
- 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 (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);
- documents that are too large to be scanned as a single full-sized image (photo reductions may be made which are then scanned, and/or multiple scanned images may be captured from the original).

Photocopies should be examined before scanning to ensure that there is no significant loss of information (see 5.2.1) compared with the original document.

Where photo reductions are made, checks should be made to ensure that there is no significant loss of detail (see 5.2.1) in the scanned images compared with the original, caused by the effective resolution of the image (compared with the original) being reduced.

Where multiple images are captured from large documents, the images should be overlapped to ensure that there is no significant loss of information (see 5.2.1) at the edges between adjoining images.

EXAMPLE

A stamp containing the word 'PHOTOCOPY' was purchased and placed near each photocopier. Procedures were introduced to ensure that all photocopies made were stamped as soon as the photocopies were made, and checked for good quality.

KEY ISSUES

> There should be documented procedures for the photocopying of paper documents prior to the scanning process.

> Where photocopying has been done, subsequent users should be made aware of this fact.

> There should be documented procedures for checking the quality of photocopies.

> Where a large document is scanned as multiple images, there should be documented procedures to ensure that the whole document is captured.

5.2.3.11 Fax input

Faxes can be received by an organization either in paper format (using a conventional fax machine) or in electronic form (via a fax server hardware/software solution). The effective scanning resolution of a fax will be dependent on the capabilities of both the sending and the receiving fax systems. Effective resolution will normally be either the lowest common denominator between the two systems or lower, depending on the settings of either system. This can dramatically affect image quality.

When paper faxes are involved, all the issues regarding paper originals in BS 10008 and the Code need to be taken into account when developing and implementing scanning systems.

Where fax input is entirely electronic, then the issues regarding data file input should be taken into account (see 5.2.6). In some applications, where the origin of the fax input is important, security measures should be introduced to check the source.

It may be important, at a later date, to be able to ascertain whether a particular document was received by fax. Thus, a suitable indicator of this fact should be created and stored.

KEY ISSUES

> Where paper faxes are input into an information management system, take into account original paper document handling procedures.

> Where electronic faxes can be received, check their authenticity prior to them being imported into the system.

> Keep a record of which documents were received by fax.

5.2.4 Image processing

5.2.4.1 General

Image processing techniques are often used to improve the quality of an image. Their use should be carefully controlled, as they can affect the quality of the image produced, and therefore the evidential weight of the stored images. With many modern scanning systems, a significant amount of image processing may be automated within the scanner. Where operator controlled image processing facilities are available for use within the scanning system, details of which facilities are to be used for a particular document should be stored and made available to users on request.

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 and used either on a scanner workstation or on a separate device.

The more common techniques include:

- de-skew;
- despeckle/background clean-up;
- black border removal;
- form removal.

To provide optimum image output, or to improve recognition rates for an automated data capture process, post-scanning processes may be performed.

Image processing facilities should only be used with care, as they can have a negative effect on the information content of an image. Any processing performed on the digitized image should not affect the integrity of the image as a true facsimile of the original. For each different process, the effect on the image should be recorded. This check should be performed on a sample set of documents, which should be scanned with the image processing turned on. Resulting images should be compared with the original documents, and any differences documented.

The effect of processing performed on a grey scale image prior to conversion to a black-and-white image should be checked to ensure that information content is not compromised.

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

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.

KEY ISSUES

> There should be a description of how image processing affects information content.

> For each different image processing process that may be used, the effect on an image should be recorded.

> Image processing techniques can affect evidential weight.

5.2.4.2 Document skew

Document skew is a term used to describe the phenomenon of poor document alignment (rotation) during the scanning processes. In its most pronounced form, images can appear on a viewing screen as crooked or slanted. Even a small angle of skew is likely to affect data capture processes and thus reduce data recognition rates.

Passing images through de-skewing processes may correct this problem.

5.2.4.3 Speckle, noise and background marks

Random black marks (speckles) that appear on an image may have been generated during the scanning process or may be present on the original document. These speckles may be removed by systems involving special algorithms. These algorithms assume that small isolated clusters of pixels contain no information, and may be deleted.

Care should be taken when scanning a poor quality original to avoid an incomplete character being mistakenly treated as speckle.

In more sophisticated processes, pixel patterns are analysed for size and presentation in a filtering process.

5.2.4.4 Speckle removal

Speckle removal should only be used with care, as it results in the elimination of single pixels or small groups of pixels from a digital image, giving 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, for example parts of already broken characters, punctuation marks or parts of fine detail in drawings.

EXAMPLE

A despeckle process is shown randomly to remove decimal points, thus altering the value of numbers within an image.

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 subsequent evidential weight of these images, if any doubt was raised about the completeness of the images.

Where speckle removal is used only on particular images, this fact may be documented in an operator log, or elsewhere in the audit trail, or by using index data for the relevant image.

KEY ISSUE

> Use the despeckle process with care, and only after extensive tests, to ensure image integrity.

5.2.4.5 Black border removal

When scanning documents of mixed sizes using certain scanner types (such as rotary scanners), black borders may be left around the edges of smaller documents. Black border removal entails the deletion of these large areas of black pixels.

5.2.4.6 Physical forms

The scanning of textual information on a pre-printed form is common when automated data capture processes such as OCR and optical mark recognition (OMR) replace a large keyboarding operation. To increase the accuracy of the recognition rate, images can be passed through a post-scanning process that will remove boxes, lines and pre-printed text. System set-up procedures are used to establish the size and position of boxes and lines that are intended to be removed.

In other applications, these forms are held separately prior to the construction of an image of an original document. Where these techniques are used, the form file should be controlled as if it was part of the data file.

In many applications, form design varies with time, as forms are improved and modified. Where this is the case, form designs should be kept of all versions of the form relevant to the data files being stored.

Where form removal software is used, a record should be made that the resulting image (the 'stripped form') has been the subject of form removal and an identifier of the template used for that removal should also be kept. This information should be stored in conjunction with the resulting image. A copy of the template should also be stored.

A facsimile made by merging the template with the 'stripped' form may not be a true facsimile of the original, although it may be sufficiently accurate for application use.

It may be appropriate to retain true facsimiles of the original forms, by retaining the originals, making a microfilm copy or retaining a complete image of the form.

KEY ISSUE

> Ensure that the relevant version of a form is used when reconstructing a document with a 'stripped' form.

5.2.4.7 Electronic forms

Electronic equivalents of paper forms include HTML forms and other, more functionally rich types; these can be regarded as electronic forms (e-forms).

An HTML form is a section of a document containing normal content, mark-up, special elements called controls (checkboxes, radio buttons, menus, etc.), and labels on those controls. Users generally 'complete' an HTML form by modifying its controls (entering text, selecting menu items, etc.), before submitting the form to an agent for processing (e.g. to a web server or mail server, etc.).

NOTE: For further information, see www.w3.org/TR/html401/interact/forms.html

More functionally rich e-forms provide a user interface to data and services, typically through a browser-based interface. E-forms enable users to interact with enterprise applications and the back-end systems linked to them.

NOTE: For further information, see www.gartner.com/it-glossary/electronic-forms-e-forms/

These e-forms support richer and more dynamic interactions than HTML forms.

Such advanced e-form applications include:

- XML content identification;
- multiple data callouts;
- field-level validation;
- embedded process logic.

When e-forms are used, it is important that:

- information from all the relevant sections of the form is processed;
- information from all the relevant sections of the form were correctly migrated to other systems;
- processing and logic utilized in completion of the form is documented;

- information accessed from other data sources during the completion of the form is documented (this may include information in constrained value selection lists, textual guidance for form completion, etc.).

These need to be within the scope of the organizational version control as the e-form, the information captured and external data sources are likely to change during the lifetime of usage and information retention.

5.2.5 Data file capture

5.2.5.1 Authenticity

When importing data files into an information management system, the authenticity of the data file should be established. Procedures for this should be documented. Where authenticity cannot be established with sufficient assurance, all appropriate information should be stored.

KEY ISSUE

> The authenticity of data files should be established prior to (or as part of) the importing process.

5.2.5.2 Format conversion

Data files are typically imported directly into information management systems, either without format conversion or including a conversion to a 'standard' format. Details of this importation process should be documented, along with details of any format conversions undertaken.

When importing without conversion, details of the originating software, including software name and version, should be stored. This information may be stored within the data file or external to it. Where it is stored internally, details of how the information can be accessed should be stored.

When importing with conversion, the conversion process should be reviewed for potential loss or addition of information. The conversion process may, for example, lose metadata stored in the original file. This may happen when converting from an office software product to (say) a text file format.

Conversion processes may also add data, for example, to document aspects of the conversion process.

KEY ISSUE

> When importing data files into an information management system, details of the conversion processes, including descriptions of any loss or addition of information, should be documented.

5.2.5.3 Dynamic data files

There are a number of data files which store data that changes from time to time. The frequency of change will depend upon the application, and can range from constant update (e.g. databases) to occasional update (e.g. web pages).

Where dynamic data files are captured, the frequency and timing of the capture operation should be chosen so that it ensures that all appropriate evidence is stored. This could be either at set time intervals or prior to a change of data.

EXAMPLE

A major supermarket offers internet-based shopping. Its website contains details of products on sale and offer prices. Prior to any change of price or product description, a copy of the relevant web page (or part of the page) is captured along with the time of update. This allows the supermarket to demonstrate its offer details should there be a dispute.

KEY ISSUE

> Choose the frequency and timing of data capture with care, to meet the requirements of the application in question.

5.2.6 Data capture

5.2.6.1 Data extraction

Data may be extracted from documents already stored on the system (or in the process of being stored) and entered into a computer in a number of ways, including manual systems (i.e. direct keyboard entry) and automated processes (e.g. bar code and/or QR (quick response) code reading, OMR and OCR/ICR. For more information on OCR, see 5.2.6.2).

Data capture can also be semi-automated (e.g. where data captured automatically, for example by OCR, is confirmed by manual scrutiny and editing). In each case, the issue is to convey confidence that the correct data have been captured. In practice, 100 per cent accuracy in captured data using OCR or ICR technology cannot be guaranteed. As it is not uncommon for OCR technology to misread perfectly valid data or even to add non-existent data, some level of manual accuracy checking may be necessary in evidential applications.

The user has to assess the risk associated with the existence of errors. Although the accuracy of OCR/ICR engines has improved significantly in recent years and 100 per cent accuracy may often be claimed by system suppliers, it is still recommended that alternative aids to ensure the required accuracy levels are implemented.

The same procedures for checking accuracy should be used when bar code, QR code or OMR techniques are used for data entry.

The acceptable accuracy levels may vary depending on the application and the importance of each particular data item. Improving accuracy levels will inevitably lead to increased costs, which may be justifiable depending upon the application. Where image-manipulating processes are used to improve the accuracy of extracted data, their use and effect on accuracy should be documented. Image manipulation may require adding information to, or deleting information from, the scanned image.

This may not be evidentially acceptable in all cases – such manipulation might be considered as image ‘tampering’. In this case, the image before manipulation should be retained and any manipulation should be performed on another copy of the image.

Accuracy checking procedures will typically be based on random or quasi-random sampling of batches of captured data, with comparison against the original material. In some applications, it may be necessary to assess the statistical ‘confidence level’ of such checking. Batches that fail to meet the required accuracy levels should generally be reprocessed and the results checked again to ensure that the required accuracy levels are met.

In all cases where data are captured from a stored document, the original data file or scanned image should be retained.

KEY ISSUES

> Where external data are captured for entry into the system, required accuracy levels should be specified.

> Accuracy checking procedures should be used to achieve the required levels.

> Manual checking of extracted data may be required to improve accuracy levels.

5.2.6.2 Optical recognition

The fundamental principle of character recognition is the ability to differentiate the contours of a character from the background. It is therefore important to distinguish the differences between a random form, a specified form and a semi-specified form. The random form has no pre-specified shape or description. A recognition engine is required to identify the existence or nonexistence of an unspecified shape or mark within a specified boundary, and this is often referred to as OMR.

EXAMPLE of OMR

The recognition of the choice of lottery numbers from the hand-completed form: in this case, the existence of marks at pre-specified boundaries indicates the player's choice of numbers.

OCR is the process of identifying the contours of characters and matching them to that of a specified form or pre-stored bitmaps. The characters that have been extracted are compared with bitmaps that have specific characteristics such as font type and font size. This enables the process of automatically entering printed text into a computer system without the need to key the data manually. A typical use is the conversion of an image file containing printed text into its ASCII text form that can then be edited using a standard word processor. In principle, this is a two-stage process, first acquiring an image with the use of a scanner or camera, and secondly processing the image through OCR software, the result of which is an editable, searchable computer data file.

EXAMPLE of OCR

Scanning purchase invoices received and extracting the data for automated entry into an accounting system.

The term ICR is frequently used with reference to techniques where some degree of 'intelligence' is used in the conversion software, to determine, on the basis of various characteristics of the 'bitmapped' image, which characters are present. ICR is often associated with recognizing handwritten characters that can be classified as semi-specified form, since it follows no specific font. However, algorithms based on advanced neural networks can recognize patterns associated with 'typical' handwritten text. In the case of ICR, it is therefore important to use multiple aids to improve the recognition rate. A typical example is where it is often impossible to consistently distinguish between the handwritten cursive characters 'm' and 'nn'. An aid to this would be to provide individual boxes for each character, which enforces a separation of characters.

EXAMPLE of ICR

Reading the handwritten text from the National Census 2001, and converting this into machine-readable text.

Some newer forms of data capture are being developed, based on artificial intelligence (AI) and on the automated drawing of inferences, for use within information management systems. Use of such systems should follow the procedures discussed in conjunction with the older technologies detailed in this section.

5.2.6.3 Operational considerations

It is particularly important to ensure that the quality of input is optimal and is free from any unintended data or mark. For example, it is crucial to ensure that the scanned images or photocopies do not contain any mark due to the presence of dirt on the scanner or photocopier panel, especially in the areas where the recognition engine is to be used.

To perform at its best, OCR software should be presented with a good quality image free from skew, with the characters separated from each other and complete in outline. The image should be of a high enough resolution to provide sufficient detail for accurate recognition, but not so great that an excessive amount of computer power is required to handle it. As a rule of thumb, the size of the text measured in points (pt) multiplied by the resolution in dpi should equal 2,400 or greater, that is 300 dpi is generally satisfactory for type sizes down to 8 pt, while for type sizes down to 6 pt, 400 dpi is more appropriate.

Most OCR engines are based on the analysis of bitonal (or, in some cases, grey scale) data. Most engines are also optimized for 200 dpi. Therefore, most colour images are thresholded before analysis. As the font colour against the paper colour affects the OCR data quality, it is also important to keep in mind the compression ratio and font size.

Although some OCR engines have been available for a long time, the technology has become accurate and time-efficient enough to be a valuable addition for automatic indexing. However, the particular technology used needs to be assessed prior to its use to ensure that the appropriate accuracy can be achieved.

In the case of ICR, constraint boxes are used to separate characters, and users are often asked to print in upper case letters and numbers. The boxes are often called ‘drop out boxes’, printed in an appropriate colour that is automatically filtered by the scanning system. The scanning system filters the boxes out and leaves only the text entered visible to the recognition engine. ‘Registration’ and ‘anchor points’ are used to align and identify forms. In some cases, these are also used to de-skew the forms. The identified fields often have some constraints added to interpret the correct character. Depending on the software used, it is also possible to attach dictionaries to improve accuracy. Also, some field inputs are cross-verified against data from other systems to ascertain data validity.

Most scanners on the market today are more than capable of producing an image of usable quality for the purpose of OCR, but there are some points that should be considered in scanner selection.

- If it is intended to scan and use OCR on large volumes of text, then a scanner with an ADF – a multi-page sheet feeder – might be appropriate.
- If the OCR volume is low, then a hand scanner might be sufficient, but it should be noted that it takes a steady, practised hand to give an even image. An A4 page will often take two or three passes which will require ‘stitching’ together before recognition can take place.
- There are many smaller sheet feed scanners appearing on the market that are ideally suited to OCR use. They provide very acceptable image quality, and are small. Often they are equipped with a 10- or 20-page sheet feeder for low to medium volume jobs.

- The use of colour scanners may improve the accuracy of the OCR process, particularly where it is possible to select the colour of the lamp (or the filter over the image sensor). If this is possible then it can be an advantage in situations with text on coloured backgrounds. Scanning with a red lamp or filter will 'drop out' or reduce a red or pink background to white, giving a better differential between the text and its background. These results may also be achieved by the use of image processing systems as part of the scanning software.

The result of an OCR process is unlikely to be 100 per cent accurate. Therefore, it is appropriate to use this technology for capturing index data only, in applications that use multiple index keys. The use of this technology is becoming common as a data capture method for 'free text retrieval' systems.

The text file that results from the OCR process output is an unformatted textual copy of the original. In terms of its weight as legal evidence, it is most likely to be inadmissible. This technology is therefore appropriate only as a method for facilitating the search of documents.

Where OCR technology is used to assist indexing and searching, the image input file should be the entity that is retrieved and used as 'the document'. Any text file produced is likely to be of reduced evidential value compared with that of the original image file. This caveat will apply even to situations in which the conversion software leaves parts of the output as images, where the software cannot identify with confidence what the characters should be. In such cases, there may still be some doubt about the accuracy of those parts of the document that have been converted to character (as opposed to image) format.

Where evidential value is important, it is advisable to keep digitized images of documents, rather than to assume that versions converted to full or partial text will be satisfactory when presented in a court of law. This does not mean that such versions should not be used within an application, only that the original scanned images should also be preserved.

5.2.7 Metadata capture

DEFINITION

Metadata – data about data

For example, metadata may be information about context and/or the relationship with other data. Metadata may render the data understandable and meaningful

When data files and documents are created or imported, care should be taken to ensure that all the relevant metadata are also transferred. Care should be taken to ensure that all necessary metadata are captured, to ensure that the data files and documents have the correct interpretation placed on them.

The content of metadata information may need to be reviewed for completeness and appropriateness. The availability of a full metadata set, with an appropriate content, will increase the evidential value of the information to which it pertains. The use of an appropriate metadata schema should be considered.

EXAMPLE

A definition of 'payment due date' could be a specified number of days after an order is placed, as opposed to a specified number of days after an invoice is issued.

KEY ISSUES

> Metadata content should be appropriate within the context of the information to which it relates, within the application being processed.

> There should be documented procedures designed to ensure and prove that all relevant metadata are captured at the appropriate time.

5.3 Self-modifying files

5.3.1 Data files

Some data files, particularly those generated by word processor, presentation or spreadsheet programs, may contain automatically executable code (often referred to as 'macros'), which can have the effect of modifying the file each time it is retrieved, viewed or printed out.

EXAMPLE

A word-processed file containing a 'current date' field is self-modifying. Whenever the document is displayed and/or printed, 'today's date' will be displayed, thus the evidence of the date on the document that was sent out of the organization will be lost.

The existence of such code within a file means that the file cannot be frozen in the sense used in BS 10008 and the Code. Each time the file is retrieved, it may appear to be different, even if the stored file has not been changed in any way. This is potentially problematic, in that it may be very difficult to determine what changes are actually being made in the retrieved file or what the original data file contained.

It may be difficult to assess what evidential weight will be attached to such a file. It is also possible that an action may be taken on the basis of the information displayed when a file is retrieved on one occasion, but that information could appear to be different on another occasion.

Where the use of such facilities is unavoidable, procedures need to be in place to ensure that authentic copies of the original information can be produced.

Where a data file that contains self-modifying code could be distributed to other parties, or stored for archival purposes, a mechanism that removes the code should be implemented before distribution or archiving.

KEY ISSUES

> Take care with files containing self-modifying code.

> Wherever possible, avoid using these facilities.

5.3.2 Databases

The equivalent use of macros on structured information held in databases is less common than for executables within unstructured or semi-structured information like spreadsheets, presentations etc. However, since a structured database can contain any of these file types care should be exercised in the event of these being present.

It is more usual within databases to create new field information content as a result of an automated procedure than to simply change the visualization at the point of rendering as with the executable in the spreadsheet, presentation etc. As such, it is more appropriately addressed under the category of version control (see 5.7.1.3) rather than as a self-modification.

Care should be taken to ensure that the procedures, either self-modification or version controls, are clear, understood, documented and applied.

5.4 Compound documents

5.4.1 Compound data files

Information may be constructed from a number of separate, stored data files, to create a compound data file.

> EXAMPLES
>
> A word processing document including a linked (i.e. not incorporated) spreadsheet.
>
> Linked voice recording and video clip.

The important issue in these cases is to be able to reconstruct a specific piece of compound information at some time in the future. Such a reconstruction will require knowledge of the status of all parts of the compound document at any given point in its history.

It is important that all parts of a compound data file are stored in compliance with BS 10008 and the Code.

KEY ISSUE

> Where compound data files are used, audit trails should be such that the historical content of the data file can be assessed at any relevant time.

5.4.2 Compound image files

Where an image is captured as an individual file on a scanning system, and parts are electronically separated to be processed in different ways, these parts should be stored after processing, along with information identifying their respective locations within the original image, to allow accurate and unambiguous reconstitution of an authentic facsimile of the complete image. This separation into parts may be done under operator control (i.e. on screen) or via software designed to separate out parts of an image.

> EXAMPLE
>
> An image contains photographic material and text. These are separated by an operator. The photographic part is kept as a full colour image, while the text part is stored as a black-and-white image.

KEY ISSUE

> There should be a description of how compound image documents are managed by the system, which could be achieved by linking shared metadata related to the relevant transaction.

5.4.3 Hybrid compound documents

Where a compound document consists of one or more data files and one or more paper files, a combination of the guidance from 5.4.1 and 5.4.2 needs to be implemented.

EXAMPLE

A health trust holds its patient records as a combination of records in digital format (database records, word processed letters etc.) and paper/film files (paper based records, paper based charts, X-rays etc.). Care is taken to ensure that the appropriate physical records are appropriately matched (and referenced) by the digital systems.

5.5 Information in structured databases

5.5.1 Introduction

Application systems and information are the transactional lifeblood of most organizations. This information is typically managed in structured databases and is at the centre of the application systems that operationally power the organization on a day-to-day basis. The range of these systems managing information in structured databases is wide and includes, amongst others, enterprise resource planning (ERP), finance and accounting, human resources, customer relationship management (CRM) etc.

5.5.2 Similarities of structured and unstructured information

Whilst many industry commentators have observed that the unstructured, document centric, information in an organization may be four or five times larger than the structured information, the importance to the organization and the likelihood that this information will need to be used in dispute resolution cannot be underestimated.

In order that structured information held in databases can be relied upon in the event of disputes there are many similarities with the integrity and authenticity aspects of unstructured information.

The policy, duty of care, processes and procedures, enabling technology and audit of the management system are equally important.

Where information stored in structured databases is within the scope of the information management system, the policies, processes, procedures, enabling technology and audit should each specifically reference this information.

5.5.3 Key areas

5.5.3.1 General

In spite of many similarities, there are a number of areas that need to be considered specifically as they are not directly analogous to the integrity and authenticity issues of unstructured, document-centric, information.

5.5.3.2 Extract, Transform and Load

Extract, Transform and Load (ETL) refers to a process in database usage and especially in data warehousing that involves:

- extracting data from outside sources:
  - an intrinsic part of the extraction involves the parsing of extracted data, resulting in a check of whether the data meets an expected pattern or structure. If not, the data may be rejected entirely or in part. If information is rejected without adequate reason and record then the integrity or authenticity of the complete data set may be compromised.
- transforming it to fit operational needs, which can result in significant changes to the information. Any such change, from any of the following actions, should be able to be explained and justified:
  - o selecting only certain columns to load (or selecting null columns not to load). For example, if the source data has three columns, then the extraction may take only two of them and not the third. Similarly, the extraction mechanism may ignore all those records where one field is not present or zero;
  - o translating coded values (e.g. if the source system stores 1 = male and 2 = female, but the target is M = male and F = female);
  - o encoding free-form values (e.g. mapping 'Male' to 'M');
  - o deriving a new calculated value (e.g. sale_amount = qty * unit_price);
  - o joining data from multiple sources and de-duplicating the data;
  - o aggregation (e.g. summarizing multiple rows of data);
  - o transposing or pivoting (turning multiple columns into multiple rows or vice versa);
  - o splitting a column into multiple columns (e.g. converting a list, specified as a string in one column, into individual values in different columns);
  - o disaggregation of repeating columns into a separate detail table (e.g. moving a series of addresses in one record into single addresses in a set of records in a linked address table);
  - o lookup and validate the relevant data from tables or referential files for slowly changing dimensions;
  - o applying any form of simple or complex data validation. If validation fails it may result in a full, partial or no rejection of the data and so none, some or all the data is handed over to the next step, depending on the rule design and exception handling.
  - many of the above transformations may result in exceptions, for example, when a code translation parses an unknown code in the extracted data;
  - it is not particularly common to have a complete audit trail of the changes made to information during operational transform. Frequently it is only test data through the transform process that will be rigorously checked; if this is the case then the test plan and test results should be retained to be able to justify or explain information changes in operational use;
- loading it into the end target (another database, data mart or data warehouse).

It is important to note that ETL processes can involve considerable complexity, and significant operational problems can occur with improperly designed ETL systems.

Where ETL has been deployed, all processes should be fully documented, as the source and target databases may have, as a consequence of the ETL process, what appears to be the same information that is not the same in the different databases involved.

Where ETL techniques are employed for information stored in structured databases within the scope of the information management system then:

- changes or loss of information caused by extraction, translation or load should be evaluated and accepted by the organization;
- processes and acceptance criteria should be documented;
- test plans, scripts and results should be retained;
- where audit trails of ETL operations are generated, these should be retained for as long as the information itself.

5.5.3.3 Database schema

For a database, the schema defines the tables, fields and relationships of the database itself. This gives context to the data held in these tables and fields enabling it to be information.

A field in a database may contain a number which could be a quantity of an item, a telephone number or a product code; the schema should describe which of these (or others) the number actually means. In this respect the schema can be regarded as metadata that gives context to the field enabling it to be regarded as information rather than simply data.

Sometimes organizations will use a field for a different purpose to that intended and this can be a cause of confusion if not properly considered and documented. In this event, the contents of a field and the schema expectation may be completely different. Such usage is often because the organization does not wish to undertake a costly bespoke modification to a packaged application system. This type of field misuse should be avoided; if it is employed it should be fully considered, approved and documented.

Another area of concern, especially with packaged application systems, is bespoke extension of the standardized schemas. This will frequently generate customized tables and fields with organization specific relationships to the standard application system. All such custom schema and application components should be properly considered, justified and documented. This type of customization is frequently the cause of considerable cost and difficulty when the standard application is upgraded to a newer version and can result in compromise to the integrity or authenticity of the information within the database.

Special care needs to be taken when any schema is modified that the changes are justified and documented and that data migration during the change process is properly tested and results recorded.

When a schema is changed, it is quite common for additional checks to the contents of fields to be applied. However, in practice, the effort to retrospectively apply these additional checks to historic data can lead to a situation where the migrated data is not validated against the new rules but migrated 'as is' when the new schema version is introduced. This can lead to situations where old data is in a different format or does not match the validation rules of newer data in the same fields. Such situations need to be able to be explained satisfactorily in the event of a challenge to information authenticity or integrity.

5.5.3.4 CRUD

A common approach to four basic functions implemented for information in DBMS are covered by the term 'CRUD'; this acronym expands to:

- Create;
- Read;
- Update;
- Delete.

There is commonality of the issues surrounding creation, reading and destruction between unstructured and structured information. However, updating is significantly different for structured information in databases from unstructured information.

When unstructured information is updated, a new version is created (see 5.7 on version control) and the information management policy will indicate how superseded versions should be managed. In databases, the update will change the contents of a particular piece of data in the database.

The transaction that resulted in the update to a particular field *in situ* may or may not be retained and this is an area that should be considered and the results of that consideration documented.

If the updating transaction and the content before the update are not retained then it may be necessary that the 'before' and 'after' contents of the field are able to be re-created in the event of a challenge to integrity or authenticity; this could be within the audit trail.

The audit trail and/or the definition of the process/procedure should indicate who or what was responsible for the creation, updating, reading or deletion of information in the database. It is worth noting that this responsibility could be a person, a device or an application component/service.

5.5.3.5 Master Data Management

A challenge facing many organizations is master data management (MDM) or a lack of it.

MDM aspires to ensure there is a 'single version of the truth' across the systems used by an organization as opposed to multiple, inconsistent versions of the same thing held in silo'ed separate systems.

> EXAMPLE
>
> Customer records are held separately in a financial system, a CRM system, and a service or warranty system. The customer details are updated separately in each system and they often get out of sync with each other.

MDM aims to avoid the need for separate updating of systems by keeping a single master that is used by the different systems and is, as a result, consistent.

Where MDM is deployed the responsibilities for the master need to be clearly documented.

Where MDM is not, or only partially, deployed, the different versions of the same information should be understood and documented so that a challenge attempting to discredit on the basis of differences between what appears to be the same information can be rebutted.

5.5.3.6 ACID

ACID (Atomicity, Consistency, Isolation, Durability) are properties that ensure database transactions are processed reliably; the long established relational database management systems (RDBMS) achieve these automatically but it should be noted that a number of NoSQL databases that have been introduced to meet the demands of big data do not include ACID transaction support.

A DBMS that does not include ACID transaction support might not be updated with every transaction posted to it; this may not be a problem for an organization using such a DBMS (e.g. in a high velocity data gathering situation the fact that an individual transaction from a sensor reading the temperature of a component every second might be lost because the DBMS did not confirm the update process may not be critical since the next reading is only one second later).

If the database is not designed to be ACID then the reliability and trustworthiness of the information content in the database may be questionable.

If the DBMS deployed is not ACID then the organization should evaluate the impact and formally record evaluations and decisions. It should be noted that suitable application design may meet the ACID transaction support criteria even if the DBMS does not.

5.5.3.7 Data quality

Data is of good quality if it is fit for purpose and correctly represents the real world construct they refer to. Aspects of data quality include:

- accuracy;
- completeness;
- consistency;
- integrity;
- reliability;
- uniformity;
- validity.

Data, as held in databases, will often suffer data quality degradation over time because it normally represents a 'point in time' record that was validated at capture but has not been re-validated subsequently.

Maintaining data quality requires going through the data periodically and cleansing it. Cleansing is normally the detection and removal of anomalies or duplicates through a workflow process.

Care needs to be exercised to verify correctness of the results of cleansing as it naturally means data has been either changed or discarded. Therefore any process that addresses and improves data quality should be properly documented and audited.

5.5.3.8 Transaction records vs. updated fields

Is there a need to be able to show, historically, what the database was at a specific point in time or not?

In many systems utilizing a DBMS there are requirements to be able to show what the information contents of a field were before and after an update or deletion, to avoid or resolve a dispute, whilst in other situations this may be superfluous.

The organization should evaluate and record decisions taken in this regard.

When the before and after update information is needed then this may be achieved by either retaining the before and after update content or by retaining the transaction and being able to deduce the database state preceding the update (this latter procedure may not be feasible for particular updates).

Taking this to the next logical stage, there may be circumstances in which it is necessary to be able to show what the information state of the whole database was at a particular point in time, rather than simply the contents of a particular field or table. This is most likely to be in situations where information from the database is routinely required to be used as evidence.

5.6 Big data considerations

5.6.1 Definition

Key point – just what is 'big data'?

Big data is a term that can have multiple interpretations as to meaning; there are about as many definitions as there are players promoting benefits of adopting it as an approach. There are a few common aspects of the varying definitions; amongst the common aspects are '3 Vs' – Volume, Velocity and Variety.

The first, Volume, is about scale and is inextricably tied to the keyword 'big', Velocity normally means that the data is both captured and needs to be accessed/analysed rapidly and Variety frequently means that the data is not consistently structured.

In many situations referred to in the context of big data the variety is very wide and the content is quite unstructured (e.g. social media messages), yet in others the content is highly structured even if there is a large variety of structures (e.g. smart meter readings, process or incident event logs etc.).

However, there is a growing recognition that these three attributes are not sufficient. There is a clue to the fact that there needs to be a wider definition from the Roman numeral 'V': there should be a group of five 'Vs' rather than three. The two additional 'Vs' being Value and Veracity.

5.6.2 Value and Veracity of 'big data'

The first three terms, Volume, Velocity and Variety, do not significantly impact evidential weight considerations; however, there is little point in an organization capturing or retaining big (or, for that matter, small) data unless it has Value that is recognized by the organization.

If the Veracity of data are questionable then it's Value, and the value of any interpretation of it, will be reduced. The value of big data is dependent on its veracity; therefore, it is important to protect the evidential value from actual or potential compromise.

5.6.3 Actions resulting from 'big data'

Decision making with big data will be based on the analysis, interpretation and inference using the data; and consequently it is that use that the big data is put to, rather than the data itself, that is the reason to capture and retain it.

Where analysis is performed then the output of the analysis should be treated as new information and the criteria by which it was extracted from the original data should be retained with the output itself. Otherwise the output, and any subsequent actions based on it, will be subject to justifiable critique.

5.6.4 Key difference

Big data will normally be retained in a database; however, there are significant differences between the management of big data and the information in 'line of business' structured databases.

It is extremely common for some of the data in structured databases to be updated as a result of a specific transaction or process.

This is not the case with big data implementations which normally contain the large volume of transactions as the big data itself. These transactions are rapidly ingested into the management system but are not subsequently updated *in situ*. Consequently, there are many similarities regarding the authenticity and integrity of these big data transactions with documents (as described elsewhere in this publication); in fact, they could realistically be regarded as specific documents (using the full scope of that term as used in this publication).

Where big data is within the scope of the information management system then:

- policies, processes, procedures, enabling technology and audit should each specifically reference this information;
- information generated from the big data should be addressed separately from the big data within the information management system.

5.6.5 Retention of 'big data'

Like any other information big data should be retained for no longer than is necessary; therefore, it should be included within the organization's information management policy and retention schedule.

An organization may have multiple sets of big data and each may have different information management policy and retention attributes.

5.6.6 'Big data' from third parties

It is increasingly common for data sets categorized as big data to include information proprietary to the using organization combined with data sets owned and provided by third parties.

A typical example of this is the use of data that is 'overlaid' onto mapping data to aid understanding and interpretation. The map is normally not owned or licensed and hosted by the organization but is a third-party data set, often provided as a cloud service.

Where big data includes third-party data sets, it may be necessary to access each third-party data set as it was at a particular point in time to be able to reproduce the complete data set for the purposes of dispute resolution.

Long-term access to a point in time version of a third-party data set may be outside the scope of the service provider's terms and conditions and, in this case, the organization should implement a plan to accommodate the non-availability of that data set.

5.7 Version control

5.7.1 Information

5.7.1.1 Unstructured and semi-structured documents

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 to maintain the link between the versions. There should also be a procedure for authorizing and implementing new versions.

In some applications, it might be necessary in the future to access particular versions or to be able to trace the revisions of the document. In these applications, each version should be treated as a new 'original' document for the purposes of BS 10008 and the Code.

It is important to realise that different versions of a document may have different retention and disposal schedules; for example, final, approved versions may need to be retained for longer than intermediate drafts.

> INFORMATION – Websites
>
> Many organizations need to be able to evidence the state of their website at the time of a particular transaction or information request; this is a form of version control.
>
> Version control of individual documents and of database schemas and information fields have been addressed; in this context, social media messages, emails, etc. are addressed under the wider classification of documents within the specification (BS 10008).
>
> However, it is important to recognize that the information content and representation of websites may have different evidential usage and retention requirements from the individual documents accessible on that website; this is frequently as a result of the dynamic nature of a website meaning different versions may be generated with increasing frequency, even dynamically as contents of databases change.

Additionally, there may also be differing version control processes and procedures from those applied to the documents published on the website or of the information in databases accessed via the website.

Consequently, organizations need to consider inclusion of their internal and external facing websites within their information management policy and associated procedures, audit, etc. and consider whether the organization's version control procedures need to recognize the differing evidential requirements of these websites.

An aspect of such procedures is that the visible representation of a website may differ dramatically depending on the network and the hardware and software capabilities of the device accessing the website.

KEY ISSUE

> Where various versions of a document can exist, mechanisms for version control (including retaining access to previous versions where necessary) need to be implemented.

5.7.1.2 Information content of structured databases

As has been discussed elsewhere (see 5.5.3.4), structured database content is frequently updated *in situ*.

Where this is routine it is outside the scope of version control but, as noted previously, there may be circumstances where it is important to be able to ascertain the contents of a particular record or field before and after update.

5.7.1.3 Databases

5.7.1.3.1 General

Whilst unstructured and semi-structured information will have significant structure (e.g. file type) and metadata associated or contained with the information itself this tends not to be the case with structured databases.

Therefore particular attention needs to be applied to version control of the database schema and the DBMS itself.

It should be noted that suitable documentation of schema or DBMS may be dependent on third-party software or service providers.

5.7.1.3.2 Schema

When a database schema is changed there are many aspects that should be addressed in order that the schema change is not the cause of unintended compromise to database content authenticity or integrity.

All database schemas produced in order to comply with BS 10008 and the Code should be maintained under a version control system.

Previous schema versions will need to be stored for at least the same length of time as that for which the relevant database content is maintained. This will ensure that the relevant version can be identified any time in the life of the stored information.

Access to previous schema versions may be required so that, for example, if a database record or field is presented in court as evidence, the database schemas that 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 challenged.

5.7.1.3.3 DBMS

When a DBMS or DBMS version is changed there are many aspects that should be addressed in order that the DBMS or DBMS version change is not the cause of unintended compromise to database content authenticity or integrity.

In order to comply with BS 10008 and the Code all DBMS and DBMS version changes should be under a version control system. Transition test plans and results should be retained as it will not normally be realistic to retain a superseded DBMS or DBMS version for a significant time.

5.7.2 Documentation

All documentation produced in order to comply with BS 10008 and the Code needs to be maintained under a version control system.

Previous versions will need to be stored for at least the same length of time as that for which the relevant information is maintained. This will ensure that the relevant version can be identified any time in the life of the stored information.

Access to previous versions may be required so that, for example, if a data file is presented in court as evidence, the policies and procedures that 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 challenged.

> EXAMPLE
>
> Where it is not possible to be certain of the scanning procedures used to capture a particular document image that is several years old, and of 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.

KEY ISSUE

> Maintain compliance documentation under a version control system, and in accordance with the retention schedule.

5.7.3 Procedures and processes

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

KEY ISSUE

> Use a formal change control procedure for managing changes to systems and procedures.

5.8 Storage systems

5.8.1 System integrity

Facilities should be provided within the system to ensure that the integrity of data is preserved throughout the system, including during the transfer of this data to and from the storage media. Regardless of the storage medium chosen, or the system environment in which the medium is used, these procedures should detect and/or prevent modifications being made to stored information from the time of capture to the time of disposal.

5.8.2 Use of checksum

A common approach to managing system integrity is to utilize a checksum calculated immediately after the data file has been captured. This technique ensures that any errors in data file transfer between subsystems may be detected automatically and with certainty.

Such a method on its own does not necessarily cover all the possibilities for malicious manipulation of the data between the time of capture and the eventual time of retrieval from 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, but in this case consideration should also be given to protection of the audit trail from tampering.

KEY ISSUE

> Where appropriate, use a checksum to check for data file changes during storage.

5.8.3 Software systems

Information is typically managed by computer software. Such software may be a facility available in the operating system (such as individual or shared drives), or may be a formal electronic document management system (EDMS) or electronic document and records management system (EDRMS).

Irrespective of the software being used, it is important to retain records of the software used and how it has been configured.

Documentation should be produced and updated where necessary to ensure that a record of the software used to manage the storage of information over time is retained. This documentation should include details of the operating system and additional software where used. Details of the version numbers and configuration details should also be retained.

KEY ISSUE

> Use a formal change control procedure for managing changes to systems and procedures.

5.8.4 WORM vs. non-WORM systems

The risk of stored data being modified inadvertently or maliciously varies with the type of storage subsystem and medium. The ability to detect any such modifications also varies. For example, where write-once media are used, it is not normally possible to modify data once it has been 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.

Data stored on magnetic disk and other random access rewritable media may in principle be modified. With such media, the risk of data being modified is less to do with the medium itself than with the controls that are implemented by the storage subsystem and by the access software. The ability to alter data requires 'read-write' access. Well-designed systems have controls to prevent unauthorized read-write access. Users with 'read only' access are unable to modify the data. Thus, the system should maintain a secure record of all read-write accesses made.

In a system where there are very frequent file modifications, there may be a substantial overhead to record modifications made, but if a record is not kept it might prove impossible to detect any unauthorized alterations, whether 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 to detect than with random access media, since if the file that is modified is not the last file stored on the medium, then all following files need to be copied and rewritten. If the storage medium is 'offline', it could be tampered with more easily if an 'attacker' were able to gain access to it. The issues of physical security of the offline media and access control while it is online are important.

WORM systems are available based on optical disk, magnetic disk or tape technology that sometimes use software/firmware controls. Compared with rewritable magnetic disk technology, such as used within RAID (redundant array of independent/inexpensive disks) storage systems, WORM systems may be more costly per megabyte of storage and may have slower access times. Thus, there may be significant financial and performance benefits to non-WORM systems. These have to be related to the potential benefit of WORM storage from an evidential weight perspective, because of the ease of demonstration of authenticity.

There are a number of issues with the use of magnetic tape in all its various forms as a long-term storage media which need to be considered, for example a characteristic termed 'resistivity'. To avoid this problem, procedures such as the frequent copying of the information to other tapes is advisable.

This feature is frequently automated in modern tape silo systems.

In any instance where data files are copied from one medium to another, a successful bit-for-bit comparison should be performed before the old medium is reused or destroyed.

The BS 4783 series, *Storage, transportation and maintenance of media for use in data processing and information storage*, gives recommendations for the storage, transportation and maintenance of media used in data processing and information storage.

KEY ISSUE

> Review the advantages and disadvantages of write-once vs. rewritable storage media, and use appropriate systems accordingly.

5.8.5 Environmental considerations

All types of storage medium have a finite life. Manufacturers' guidelines should be followed to reduce the risk of losing data because of inappropriate storage.

Storage media should be checked on a regular basis, and in accordance with manufacturers' recommendations where available, to detect any degradation of the media. Media handling and storage recommendations should also be followed.

System hardware should be installed in locations that can be environmentally managed to manufacturers' recommendations. This may include some or all of the following considerations:

- temperature stability;
- humidity management;

- safeguards against power fluctuations or loss;
- security measures;
- protection from physical threats.

BS 7083:1996, *Guide to the accommodation and operating environment for information technology (IT) equipment* gives guidance on the accommodation and operating environment for IT equipment. It covers construction and accessibility, environmental conditions, electrical power requirements, and operational safety and security.

KEY ISSUE

> Hardware manufacturers' recommendations for the operational environment of all components of the system and the storage media used should be followed.

5.8.6 Write to media process

It is important to ensure that the electronic data file is protected at all times, including the time between capture and write to final storage. The data file will often be particularly vulnerable at this time. Thus, the point in the application processes at which data are written to storage should be as soon as possible.

In some systems, multiple storage media are used, and data files are transferred by the system on an automated basis. The actual storage media will mainly depend upon access requirements. Such systems are termed 'hierarchical storage systems', and include high-cost rapid access media and also low-cost slow access media (plus some intermediate types where appropriate). In these systems, the integrity of data files during transfer from medium to medium should be demonstrably unaffected (see also 5.8.7).

KEY ISSUE

> Ensure data files are protected from time of capture, by writing to final storage as soon as possible.

5.8.7 Media migration

Information may be stored for a considerable length of time and, importantly, for longer than the lifetime of the current technology. Thus, to ensure the integrity of stored information, it is important to plan from the outset that it may be subject to migration processes. Such processes may involve a change of:

- media; and/or
- computer hardware.

As a rule of thumb, a storage media migration process will occur approximately every five years.

Where information is moved from one storage device to another, as part of a data file migration process, details of the move should be stored in the audit trail.

Procedures for media migration should include methods by which it can be demonstrated that any related data (such as metadata) are also migrated, where this is necessary.

In the case of HSM systems and storage area networks (SANs), where data are routinely and automatically moved between storage devices, without user intervention, it may not be necessary to generate audit trail data on these movements of information. However, it will be necessary to demonstrate that the HSM or SAN was working normally when data were transferred. In practice, this will mean throughout the period from initial capture of the information to its final storage.

See also 5.8.6.

KEY ISSUES

> Where data will be stored for longer than five years, plan for storage media migration by using industry standard systems.

> Keep records of migration processes, to demonstrate that integrity was not compromised during the process.

5.8.8 Format conversion

5.8.8.1 General

The same issues related to media life (see 5.8.7) also apply to file format life. Where new versions of software used to create information are implemented, issues with backward (and in some cases forward) file format compatibility may occur.

Thus, to ensure the accessibility of stored information, it is important to plan from the outset that it may be subject to format conversion processes.

KEY ISSUE

> Where data will be stored for longer than 10 years, plan for storage format conversion, or guard against this requirement by using industry standard information archiving formats (see 5.8.8.3).

5.8.8.2 Conversion

Where information has been converted from one file format to another, details of the conversion should be stored in the audit trail. For example, a document created with one word processor program may be converted to another word processor format without changing the text within the document. From one perspective this might be considered not very different from copying a file, but if formatting is relevant to the information content, there is the possibility that the information content of the converted file may be considered to have changed.

When making provisions for converting data files, it is important to include all relevant metadata, including index data and audit trails. This additional data should also be migrated to the new technology without loss of integrity.

Records, including audit trails, should be kept of any conversion processes that stored data have been subjected to, to allow the integrity of the data to be demonstrated beyond any reasonable doubt at any time in the future.

The conversion of stored data from one file format to another can be a complex task, as it can involve:

- simultaneous migration to new storage media;
- change of data format;
- security of conversion processes;
- issues surrounding the deletion of information during conversion, for example, as part of an information retention policy;
- confirmation that all data that needs to be converted have been converted;
- assurance that authenticity and integrity are not compromised;
- inclusion of appropriate metadata;
- addition of metadata detailing the conversion process.

Where a data migration process is implemented, procedures and processes to manage these issues should be implemented. Where data are being received from another information management system (or part of a system) which conforms to BS 10008 and the Code, as part of a system migration process, then procedures and processes need to be established, implemented and documented for this process.

Some migration processes may need a format change, to enable the data to be accepted into the receiving system.

KEY ISSUES

> Migration processes should be implemented with care, to ensure that evidential value is not compromised.

> Where data files and any associated metadata are being received from a system that conforms to BS 10008 and the Code as part of a migration process, the migration procedures and processes should be documented.

5.8.8.3 Standard formats

A reliable methodology for dealing with software life is to ensure that data files are stored in an industry standard format, or that viewers for each stored format are maintained. It is also recommended that a restricted number of formats are used for long-term storage, to reduce future format migration issues.

One such file format is that defined by BS ISO 19005-1:2005, *Document management — Electronic document file format for long-term preservation — Part 1: Use of PDF 1.4 (PDF/A- 1).*[7]

KEY ISSUE

> The use of standard file formats protects against long-term access issues.

5.8.8.4 Alternative technology

Where long-term storage (i.e. greater than 10 years) is required, and where access levels are very low (i.e. less than 1 per cent), the use of microfilm storage should be considered, utilizing its technology-independent functionality to eliminate the need for migration processes.

KEY ISSUE

> Microfilm should not be discounted where very long-term storage is required, and access requirements are negligible.

5.9 Information retention and disposal

5.9.1 Information retention

Where original documents are scanned and the policy document states that it is general policy to destroy a specific type of original document after scanning, there are some instances in which an exception applies and the original document should be retained.

[7] NOTE: The PDF/A file format will evolve over time, specified by future parts of BS ISO 19005. This standard should be used in conjunction with BS ISO 32000-1, *Document management — Portable document format — Part 1: PDF 1.7.*

Circumstances where this may be required include:

- where the original document is of poor quality, so that a legible image cannot be obtained;
- where the original document contains physical amendments or annotations that cannot be identified as such on the scanned image;
- where fraud or other misdemeanour is suspected or has been identified or where litigation is envisaged or is pending.

Where the original document is of poor quality, it may be necessary to keep the original, 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 original document. Alternatively, a note may be stored which states that the original document was of poor quality and which includes details of any visible information that needs to be stored.

Where an original 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 for legal admissibility.

Procedures for the identification of information for which fraud or other misdemeanour is suspected or has been identified, or for which litigation is envisaged or is pending, should be documented. Such procedures should include the suspension of document destruction policies for this information, as destruction in these circumstances is a criminal offence.

KEY ISSUES

> There should be documented procedures that identify specific original documents that need to be retained after scanning.

> Do not destroy any documents where fraud or other misdemeanour is suspected or has been identified, or litigation is envisaged or is pending.

5.9.2 Information disposal

Procedures for the disposal of information (which may involve destruction of the information or the transfer to archival storage) at the end of the retention period should be documented. These procedures should ensure that all copies of the information are disposed of, including those held on backup systems or as 'personal' copies.

Disposal procedures should incorporate security precautions appropriate to the sensitivity of the information being disposed of or destroyed.

Where original documents are destroyed after scanning, they should be kept for at least as long as is necessary to ensure that their electronic files have been successfully written to storage and the appropriate backup procedures have been completed.

Disposal processes should be auditable, such that (for example) the disposal of a particular document can be proven. It is also important that any necessary authorization for such processes be obtained before disposal.

Where data are stored on WORM media, disposal of specific information is not possible (unless a controlled process of selective copying to new media is implemented). In some applications, it may be accepted that removal of all index references to the information being disposed of is, in practical terms, disposal of the information itself. Organizations need to check that this procedure is acceptable.

Care should be taken when deleting information on magnetic media. Typically, simple 'file deletion' actions can be reversed, and so may not be appropriate in some applications. In these circumstances, the use of degaussers and/or shredders should be considered.

When positive removal of information from the system is required, identification and deletion of all copies of the information (including in the backup media) will ensure that the necessary action is taken.

Where expungement (e.g. removal without any trace of the information ever existing) is required, procedures for the disposal of the information should be enhanced to remove any trace that the original information existed. The replacement of the information with a document stating 'expunged information' may be sufficient.

KEY ISSUES

> There should be documented procedures for the secure destruction, or transfer to archive, of stored information.

> These procedures should deal with all copies of a particular document.

5.9.3 Legal holds

Standard disposal of information in accordance with the schedules needs to be suspended where the information may be required as evidence. This is because this information could be called as evidence, even if its end-of-retention period has been reached.

This suspension of standard disposal is commonly known as a 'legal hold'.

In most jurisdictions, organizations have a duty to preserve relevant information when they learn, or reasonably should have learned, of litigation commencing or of a regulatory investigation. In some jurisdictions, the legal hold needs to be applied when litigation or regulatory investigation is reasonably anticipated but has yet to be formally initiated.

In order to comply with these preservation obligations, the organization will need to inform information stewards and custodians of their duty to preserve relevant information.

If this suspension of disposal is not enforced, and demonstrably so, there may be potential or actual destruction of relevant evidence or spoliation, which may be interpreted as malicious rather than inadvertent and consequently negatively affect the outcome of the litigation or even the burden of proof from being innocent until proven guilty to being required to prove innocence.

KEY ISSUE

> Failure to apply effective legal holds can prejudice a trial and disposal of information that may be required as evidence can be a criminal offence.

5.10 Information transfer

5.10.1 General

Typically data files and documents will be transferred to the information management system from outside the organization or another part of the organization by:

- intra-system electronic transmission;
- external electronic transmission;
- physical transfer (of electronic or other media).

5.10.2 Intra-system data file transfer

Intra-system information transfers are those that take place within the system as defined in Chapter 4. Intra-system data file transfers include:

- local area network transmissions or those within an organization's private network or virtual private network (VPN);
- transfer between structured databases as a result of an ETL (see 5.5.3.2) process;
- movement between storage subsystems under system control (e.g. in an HSM system, between cache and magnetic disk, within a SAN, or between network attached storage (NAS) devices);
- transfer between storage subsystems under operator control.

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

Data file transfers within an organization should be controlled by the application software, which should include processes that ensure that the integrity of data files transferred is not compromised. Digital signature technology is an example of such a process.

NOTE: This section is not applicable to the requirement for data file migration (see 5.8.7), where the media type and/or format of the data file may change for technology migration reasons.

KEY ISSUE

> Procedures should be implemented to ensure that the integrity of any data files transferred between systems is not compromised.

It should be noted that system integration using an enterprise service bus or the exploitation of MDM do not generally result in an intra-system transfer since a specific piece of information is accessed by multiple applications or services in a common database location; as such access control should be a consideration.

5.10.3 External transmission of data files

5.10.3.1 General

This section deals with data files transmitted between one system and another via external, wide area, communications systems. Such systems are external to the system described in Chapter 4. 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. For further detailed information about the maintenance of authenticity and integrity of information during external transfers, see BIP 0008-2.

Electronic messages (e.g. email, instant messaging or EDI) are assumed here to be equivalent to data files. The communications system may involve real-time transmission or deferred ('store and forward') transmission such as occurs in most email systems.

If there is a dispute over the authenticity of a received file, and the sender's procedures conform to those of BS 10008 and the Code, but the receiver's procedures do not, less reliance may possibly be placed on the contents of the receiver's file. Conversely, if the receiver's procedures conform to BS 10008 and the Code but the sender's procedures do not, less reliance may possibly be placed on the contents of the sender's file.

NOTE: Additional procedures may be adopted for confidentiality (e.g. to prevent unauthorized disclosure of the information contained within a data file) or other reasons (e.g. to check that the data file or document does not include malicious software, such as viruses).

The level of security risk being taken during an external data file transfer should be assessed, to ensure conformity to the requirements of the information security policy.

5.10.3.2 Compression techniques

Compression techniques may be used to reduce the size of data files, to reduce the amount of storage required, and to reduce the amount of network traffic, to improve system performance. Such techniques may be applied to data files by the system prior to or during storage. Information about the compression ratio achieved may be stored as part of the data file or its related index data, or via a separate log. In the case of image files stored in Tagged Image File format (TIFF), the compression method is automatically stored within the image file.

The compression technique used is typically application dependent, though some systems may have built-in compression that the organization has no alternative but to implement.

Where compression is used, the system should provide adequate facilities, preferably via automated means, to ensure that the requirements for quality control (e.g. checking of image quality after scanning, with ability to rescan if necessary; control over index data accuracy; control over data integrity or playback of audio or video recordings) of the compressed data file can be met.

Compression techniques can use various mathematical approaches, but all may usefully be classified into two classes, namely ‘lossy’ or ‘lossless’. With lossless compression, the decompressed data file is absolutely identical to the original uncompressed one – this is not the case with lossy compression.

KEY ISSUE

> Compression techniques should only be used in accordance with the policy document.

5.10.3.3 Lossy compression

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

Lossy compression may be suitable for photographic or other continuous-tone material, grey scale or coloured documents, where it can be shown that there is no significant loss of information in the scanned image. Similarly, most audio and video data file formats utilize lossy compression techniques, albeit different ones, to meet data file size constraints and thus need to be considered from an evidential perspective.

Where lossy compression is used, a sample set of decompressed data files should be compared with the originals to check that there is no significant loss of information.

Compression ratios should be chosen where possible such that no significant information is lost due to the technique used. The maximum acceptable compression ratio may be determined via the sample set of originals, and may vary between documents in the sample set. It may be necessary to decide whether to use different compression ratios for different documents or to use a single ratio for all documents. If the latter approach is adopted this will usually mean that the average image, audio or video file size will be larger but the speed of processing will be higher because of reduced operator intervention.

Where it is important that there should be no loss of information in a scanned image other than that due to the scanning resolution, lossy compression should not be used.

EXAMPLE

Lossy compression should not be used on radiographs (i.e. medical and engineering X-ray images), where no loss of detail is acceptable.

5.10.3.4 File integrity

BS 10008 and the Code are concerned with the integrity of data that has been transmitted to another party, and with the integrity of data received from another party. BS 10008 and the Code are not directly concerned with the transmission service. In order to comply with BS 10008 and the Code, it should be possible to demonstrate that a data file that was transmitted to another party has not been altered since it was transmitted. It should also be possible to demonstrate that a data file received via transmission from another party has not been altered since the time of receipt. Some of the techniques employed to demonstrate that no changes have been made to a data file in this way can also be used to demonstrate that no changes were made during the transmission.

Differences between sent and received data 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 of a paper document are equivalent. The primary need is to show which file is the source, and which file is the copy, that is which file existed first.

In some instances, this requirement can be met by comparing the times at which the two copies of the data file were stored. If system time clocks are accurate, a received file should have been stored later than that at which the source file was transmitted. Particular care needs to be taken to ensure that differences in settings of the time clocks of different systems and differences in time zones are considered and accounted for. A key issue becomes that of being able to demonstrate the reliability and accuracy of the timings of the two events.

5.10.3.5 Data added during transmission

It is common, particularly with email systems, to add additional data to data files (messages in the case of email) after they leave the sender but before receipt. This may be, for example, to add either a formal disclaimer or confirmation that the message has been checked for the absence of malicious software (e.g. viruses).

Where the possibility exists that there might be a challenge to the authenticity of a document or data file that has been modified during the transfer process, the sender should be able to reconstruct the communicated document. In order to be able to do this, the original skeleton of the document should be stored, together with information regarding the additions used for any particular transmitted document. A similar approach may be taken when correspondence is drafted from standard form components (templates) plus possible customization.

5.10.3.6 Use of digital signatures

Digital signatures (see 5.14.4) may be used to permit confirmation that a received data file is exactly the same as was sent, and where appropriate to confirm the identity of the sender. This may be an identity that has been validated by a third party. For further information, see BIP 0008-3.

5.10.3.7 Proof of delivery

Where it is important to be able to demonstrate that a data file has been delivered, the sender may require that the receiving system transmits back to the sender a confirmation of receipt, which should include the transmission identifier and the date and time of receipt. Where these procedures are followed, then the risk that a data file has been modified, or has been sent from someone other than the identified sender, is reduced.

KEY ISSUES

> Data file transfers from one device to another should be controlled by the application software.

> Incorporate a transfer checking mechanism. Have all files that were sent been received, and were they received without change?

> Include in audit trails the date and time of transmission.

> Incorporate a mechanism to ensure that files cannot be accepted from unauthorized locations.

> Encrypt files during transfer where confidentiality is an issue.

5.10.4 Physical transfer of media

When data files or documents are transferred on physical media, there may be other risks to their integrity or authenticity than with electronic transmission. For instance, the carrier may have lost or interfered with the information being transferred. It is therefore important to document how the transfer is made and how compromises to data quality or completeness are to be avoided. This may include checks carried out at the point of despatch and the point of receipt. When the check at receipt indicates that information may have been interfered with, the actions to be taken should be documented.

EXAMPLE

Many organizations use locked pouches or boxes to house documents or other media during transportation. This prevents the carrier from interfering with the contents, either maliciously or inadvertently. The box is signed for at despatch, and logged in on receipt to form part of the audit trail.

5.10.5 Business process mangement and workflow systems

Some information management systems incorporate business process management (BPM) and/or workflow capabilities. 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. Each particular process will have a related process definition life cycle, which may include:

- definition;
- development;
- implementation;
- withdrawal;
- modification.

Although BPM initially focuses on the automation of business processes with the use of information technology, it can also include human-driven processes in which human interaction takes place in series or parallel with the use of technology.

EXAMPLE

A BPM or workflow management system can assign individual steps requiring deploying human intuition or judgment to relevant people and other tasks in the workflow to a relevant automated system.

In most workflow systems, an audit trail point exists at each step in the workflow. However, for compliance with BS 10008 and the Code, audit trail information may not need to be kept for every audit trail point for the same retention period. The organization should decide which audit trail points are relevant with regard to the potential evidential importance of the data within the workflow. It may be appropriate, for example, to keep audit data where every automated or manual decision point is reached. These audit trail points should be selected for the generation of audit trail data.

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.

EXAMPLE

A leading retail fund and investment organization used to retain audit records of every step of a workflow task. These records were never deleted. The volume of workflow audit data soon exceeded the volume of information to which it related.

After a risk assessment exercise, and the recognition that all appropriate audit trail data that could be needed at a later date were stored on another system, it was decided to discard all workflow audit data after a three-month period.

A record should be created, for audit trail purposes, each time a new business process is defined or an existing definition is changed. The selected audit trail points may change as the workflow processes are changed.

The system should permit an authorized user to select and/or de-select the audit trail points for which audit trail data are generated. Any changes to workflow audit trail content should be audited.

KEY ISSUES

> Where workflow systems are implemented, operational details (such as flow diagrams, process definition classifications, process definition life cycles and change management) should be documented.

> Review the evidential value of data (databases, audit trails, etc.) held on the workflow system.

> Where workflow systems are in use, relevant steps in the workflow should be defined as audit points. Audit trail data should be generated at these points.

> Records of changes to the workflows should be kept.

5.11 Indexing and other metadata

5.11.1 General

Indexing is a vital part of the process of storing information on electronic media. The index information allows for retrieval of electronic information. Where indexing information is lost, then the stored information may also be lost, or be unable to be found without excessive effort.

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.

Search tools and techniques can be a useful adjunct to manual or automatic indexing to support retrieval of structured or unstructured information. Whilst search tools and techniques can reduce the effort required to generate metadata by identifying sets of data to which metadata can then be associated, it is important to note that where indexes and other metadata exist before a search is applied then the use of that index or metadata to refine the search will improve the search performance. Therefore, search tools and techniques should not be regarded as an alternative to indexing or information classification.

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.

KEY ISSUE

> Procedures and rules for indexing stored information should be documented.

5.11.2 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.

Indexing processes may include the detection of missing or inaccurate indexes. Manual indexing systems (e.g. keying from displayed information) will not detect inaccurate indexes unless the displayed information is checked against the originals (maybe by double keying), or there is a defined sequence of information (e.g. by sequential numbering).

KEY ISSUES

> Staff involved in manual indexing should receive specialist training, to maximize accuracy.

> Consider the use of accuracy-improving methods such as double keying.

5.11.3 Automatic indexing

Automatic indexing may be achieved by, for example, the reading of bar codes or the use of OCR/ICR techniques (see 5.2.6.2).

KEY ISSUE

> Where automatic indexing is used, procedures to check and amend inaccurate index data should be documented.

5.11.4 Index storage

In order to retain access to stored information, relevant index data should be retained for at least as long as the stored information to which it refers.

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

KEY ISSUE

> Where indexes are rebuilt, relevant procedures should be documented.

5.11.5 Index amendments

Where an index entry has been determined to be inaccurate or missing, then a procedure for its correction will need to be followed.

When an index entry has been amended, there may be an advantage in retaining details of the index content before and after the change.

Where an index entry relates to deleted or expunged information, the fact that it relates to such information should be stored.

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

KEY ISSUE

> Procedures for the amendment of indexing data should be documented.

5.11.6 Index accuracy

Index data, whether for scanned images or any other types of data, may be inaccurate. While accurate indexing will facilitate the retrieval of stored information to be disclosed in court, the evidential weight of that information may be increased if its relevance and completeness can be demonstrated from the accuracy of the relevant index data. Conversely, inaccurate index data may result in the user being unable to retrieve relevant information or, alternatively, irrelevant information may be retrieved.

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 should 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.

KEY ISSUE

> Accuracy levels should be agreed and documented.

5.12 Output

5.12.1 Authenticated output procedures

Output, typically in the form of paper or electronic copies, may need to be produced for use in court. Generally, these copies need to be authenticated as true copies of the original, to reduce the likelihood of rejection by the court.

Such procedures may, for example, require the use of standard system features for the creation of a copy, plus written confirmation by an authorized person that the output 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 output.

Authenticated output procedures can be enhanced by the inclusion of a statement such as 'created from a system in compliance with BS 10008:2014, Specification: Evidential weight and legal admissibility of electronic information'.

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, including in a court of law. 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 that 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 that obtained by 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, because of loss of detail or layout differences. These differences may be unacceptable for legal disclosure, and different retrieval facilities may be required that do not involve conversion.

KEY ISSUES

> Procedures for the creation of authenticated copies should be documented.

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

5.12.2 Authentication of copies of data files

In some applications, for example, involving contractual documents, several parties may each hold copies of a data file. In this case, a dispute could arise as to which copy is a true copy of the original.

In the case of a scanned original document, the digitized image file is a facsimile 'copy' of the original document. Subsequent copies of the image file may be distinguished by recording the dates of original storage in each case. Thus, checking a copy of any data file against earlier copies allows a trace to be created back to the earliest file and, in the case of a digitized image, to the originally captured data file.

As an example of such a procedure, a digital signature could be produced, and stored securely with a trusted third party (e.g. a digital notary). In the event of any dispute, the digital signature could then be used to demonstrate whether a file is a true copy of the original file.

KEY ISSUES

> Procedures for authenticating copies of data files should be documented.

> The use of digital signatures and/or trusted third-party authentication may be advantageous in some applications.

5.12.3 E-discovery

E-discovery is concerned with finding what may be relevant to a case and then disclosing it to the opposing party in the dispute.

E-discovery is a significant focus for the legal profession because the techniques of finding electronic information and then disclosing it are markedly different from non-electronic information and so are the costs. However, the ways in which challenges to the authenticity and integrity of electronic information can be resisted are independent of the ways in which it has been discovered and disclosed.

A key issue with e-discovery and e-disclosure is being able to justify that the information set is complete. This may be because:

- e-discovery tools have missed important, relevant information because of inadequate selection criteria;
- e-discovery tools have not had access to a comprehensive information source;
- relevant information may have been disposed of either in compliance with or outside the terms of the organizational retention and disposal schedule.

5.13 Identity

Where the identity of those involved in information capture is significant, procedures which authenticate the identity of the person, organization or other entity need to be established.

These procedures should be designed and implemented to ensure that:

- the authenticated identity is bound to the electronic information;
- the proofs used to associate physical to electronic identity are approved by the organization;
- the processes used for proof of identity are retained, along with their results.

NOTE: Proof of identity is discussed in more detail in BIP 0008-3.

5.14 Information security procedures

5.14.1 Security procedures

It is prudent to adopt relevant security guidelines that are applicable to the organization concerned. Such guidelines might exist in company policies or practice, sector-specific guidance (e.g. financial or medical), national or international standards, or as legal requirements.

In the absence of internal guidelines, published documents (see Annex G) may provide a comprehensive set of information security guidelines that can be designed to meet the organization’s needs. They might provide an adequate basis for the creation of internal 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 information security policy.

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

Where appropriate, the accommodation and operating environment for information management systems and for the storage, labelling, handling, transportation and maintenance of data storage media should be in accordance with the supplier's recommendations and/or relevant national or international standards.

The central part of the system (e.g. including the file servers and data storage) should be installed in physically and environmentally secure areas, with restricted access. These restrictions should be documented.

KEY ISSUE

> There should be documented procedures in accordance with the organization's information security policy (see 2.2.3).

5.14.2 Access rights

Access to information systems and services can either be by workers or by systems and services.

Only workers with the relevant access rights should be permitted to enter data or amend stored data. There should be an effective segregation of roles to ensure that a single person cannot complete all aspects of transaction authorization, data entry or amendment, and review. It is also important to ensure a suitably detailed level of automatic logging is applied to the process to record those involved, the activities performed, and times and dates.

System access rights should be granted only after the worker has successfully proved their competence.

EXAMPLE

A network attached scanner has automated access to a network location which enables scanned documents to be directly loaded into a workflow process. In this case it is important to identify the identity of the worker scanning the document in the audit trail.

Systems and services access should be controlled to ensure that only approved, authorized and authenticated systems and services are granted access to create, read, update or destroy stored information.

It is also important to ensure a suitably detailed level of automatic auditing is applied to the process to record the systems and services being granted and utilizing these access rights.

EXAMPLE

An application is required to update different database tables automatically and autonomously when a threshold is met as a result of cumulative transaction details. The application in question may only need access rights for when the threshold is met.

EXAMPLE

For any extraction or load during ETL (see 5.5.3.2), the ETL subsystem will need authorized access to the object or target database; similarly authorized access rights are essential for system integration using an enterprise service bus.

EXAMPLE

A system accessing MDM data (see 5.5.3.5) has not been granted access to an upgraded MDM data set and continues to access obsolete information creating inconsistent results that would not be able to be either justified or recreated in the future.

These access rights should be included within the scope of the system access control system and managed appropriately.

All access rights granted and revoked are significant information that should be retained under the scope of BS 10008 and the Code in accordance with the policies of the organization.

KEY ISSUE

> Implement a managed system access control system.

5.14.3 Encryption

Cryptographic techniques may be used to improve the security and integrity of stored data. A complete data file or a part thereof may be encrypted so that the information it contains cannot be retrieved without the use of an encryption key.

Digital signatures (see 5.14.4) consist of data (keys) which, when appended to a data file, enable the user of the data file to authenticate its originator and/or its integrity. The digital signature data can be applied and checked by the appropriate use of secret keys or private and public cryptographic keys. The use of a digital signature does not imply that the file itself has to be encrypted. In many cases the file may 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. It may be necessary to check that the identity of the purported signatory has been confirmed by another trusted party. This will typically be done by checking with the certifying authority (or its agent) that issued the digital certificate to the purported signatory. This digital certificate will normally be included as a part of the digital signature and will include the Public Key of the signatory.

Where encryption and/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. Encryption key allocation and management, and certificate management where digital signatures are used, should be included within these processes. It may be appropriate to consider the use of third-party key management and recovery or key escrow services.

The person who originally was responsible for managing the keys and certificates securely within the organization may have been redeployed, be unavailable or may no longer be employed at the organization, so procedures should be implemented to ensure the timely access to and continued availability of the keys and certificates.

In some countries, the use of encryption techniques may be restricted or illegal. In these cases, it is usual that digital signatures may be permitted.

EXAMPLE

Certifying authorities are examples of commercial trust services.

'tScheme[8] is the independent UK, industry led, self-regulatory scheme set up to create strict assessment criteria against which it will approve trust services. tScheme have recently published approval profiles designed for IdP Services that are based on non-PKI credentials, which would include PIN/Password, Chip & PIN smartcards, etc. As a result the information on the approval profile is split into two sets – those for PKI-related Services and those for IdP Services.'

KEY ISSUES

> Keep encryption keys secure and make them available only to authorized personnel.

> Check that encryption can be used legally within any appropriate countries.

5.14.4 Digital signatures and electronic signatures

DEFINITION

Electronic signatures are defined in the Electronic Communications Act 2000, s.7(2)[9] as:

'…so much of anything in electronic form as —

(a) is incorporated into or otherwise logically associated with any electronic communication or electronic data; and

(b) purports to be so incorporated or associated for the purpose of being used in establishing the authenticity of the communication or data, the integrity of the communication or data, or both'

The Electronic Communications Act 2000 deals, amongst other topics, with the use of electronic signatures and related certificates in legal proceedings.

In practice, there are two general types of electronic signature – see the following definitions.

DEFINITIONS

Digital signature – data appended to a data file that allow the recipient of the data file to authenticate the source and the integrity of the data file

Electronic signature – computer data compilation of any symbol or series of symbols executed, adopted or authorized by an individual to be the legally binding equivalent of the individual's handwritten signature

As well as being able to be used in legal proceedings, digital and electronic signatures offer the possibility of demonstrating that retrieved data is exactly what was stored, and of confirming the identity of the people and/or organizations involved with the storage.

Electronic signatures are typically created with digitizing devices. Digital signatures use private and Public Keys, frequently held within a digital certificate.

[8] www.tscheme.org/profiles/index.html

[9] www.legislation.gov.uk/ukpga/2000/7/section/7

Both electronic and digital signatures are fundamentally cryptographic checksums. If an attempt at malicious manipulation is to succeed, access to the original signature digitizing device or Private Key would be required (in addition to knowledge as to the particular cryptographic algorithms utilized).

The authorized retriever of a digitally signed data file may use the signature and the relevant Public Key to verify the identity of the original signatory and the integrity of the data file. This applies to storage, workflow or transmission, whether real-time or store-and-forward transmission systems are used. Digital signatures should be used in applications where it is important to be able to confirm the integrity of a data file and/or the identity of the signatory.

In some instances, the item signed could be part of a data file (e.g. the body of an email, but not its transmission headers) and in other circumstances the item signed could be the entire data file.

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

Signature management within the organization is a key issue. Before a key is issued, the true identity of the person prior to that individual being enrolled as a document signatory should be checked. Whilst this is typically not an issue with workers, such checks may be necessary with temporary staff and contractors. Signatures should be stored securely, and access to Private Keys and algorithms should only be allowed by authorized personnel.

Digital signatures are useful techniques for integrity and binding information to a specific entity (frequently an individual). Care does, however, need to be exercised over placing too much reliance on a digital signature without understanding the underlying technology and the risks associated with it. If access to Private Keys is insecure and the algorithms are known, it may not be possible to distinguish between a valid signature and a malicious one, as there can be nothing to distinguish them. Rigorous security measures are needed in order to prevent a valid but false digital signature being generated. Also, where long-term storage of data files is envisaged, computer technology available in the future may be able to compromise current digital signatures without detection.

It should be noted that different signature algorithms and key lengths can have different strengths. Organizations should consider the options and select algorithms and key lengths that have strengths appropriate to the particular document or data file.

If a query is raised about the authenticity of a data file, signatures may be used as evidence in demonstrating that any data file stored or received by transmission contains the same information as the original data file.

Where appropriate, systems should use electronic/digital signatures and/or security copies that are stored in different locations, possibly involving trusted third parties (see 5.16.9), as integrity maintenance methods.

For further information on electronic identity management, see BIP 0008-3.

KEY ISSUES

> Electronic and/or digital signatures should be used where it is important to be able to confirm the integrity of a stored or transmitted data file and, where appropriate, the identity of the party concerned.

> There should be secure processes for the generation and issue of digital keys.

5.14.5 Other authentication techniques

As well as digital signatures, as discussed in 5.14.4, which are more closely associated with asymmetric cryptographic techniques, there are other authentication technologies being developed, typically based on symmetric techniques. Examples of such technologies are message authentication codes (MACs).

Where these technologies are used, procedures that ensure their appropriate use should be developed. Reference to the controls recommended for digital signatures should be made as a guide to the controls recommended for these technologies.

KEY ISSUE

> Where other authentication techniques are used, procedures for their appropriate use should be implemented.

5.14.6 Email

Digital signatures are often added to email messages as a security measure.

Where a digital signature is added to an email, the signature algorithms should typically be applied to the message body, including the message itself and any attachments appended. Additional information may be added to the message (e.g. the organization's disclaimer). This will not be within the signed message. However, the recipient will still be able to access and confirm the digital signature and so ensure that the message has not been tampered with during transit.

NOTE: There are standards on how and where digital signatures are to be applied to email messages. Where interoperability is a requirement, the use of recognized standards such as the s-MIME v3.1 specification is of value.[10] These standards can break email messages into component groups, thus allowing the signatures to be applied in specific ways to meet different requirements.

Where email archives are used, it is key to be able to demonstrate that the whole of the message has been stored and retrieved without change. In this case the digital signature should encompass the whole of the message.

KEY ISSUE

> Digital signatures should be used as an email security measure where appropriate.

5.14.7 Malicious software

To protect stored data 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.

In systems that 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 that have been frozen have not been altered or removed.

[10] Secure/Multipurpose Internet Mail Extensions (S/MIME) Version 3.1 Message Specification. Available at: www.ietf.org/rfc/rfc3851.txt

EXAMPLE

Malware, short for malicious software, is software used to disrupt computer operation, gather sensitive information, or gain access to private computer systems.

Malware includes computer viruses, ransomware, worms, Trojan Horses, rootkits, keyloggers, dialers, spyware, adware, malicious Browser Helper Objects (BHOs), rogue security software and other malicious programs. Detection and recovery from malware attack and contamination with associated information destruction or modification can be very time consuming and expensive.

KEY ISSUES

> Review the possibility of accidental or deliberate modification to data files during storage, and put mechanisms in place to detect and/or rectify these issues.

> Use anti-virus (and other) systems to prevent malicious alteration of stored data files.

> Use power failure protection mechanisms to prevent corruption of stored data files.

5.14.8 Backup and system recovery

Effective procedures for the backup of electronic files provide sufficient up-to-date copies of data to be used in the event of loss or corruption of part or all of the 'live' data. It is vital that backup data include all associated information (such as index files and audit trails), so that a complete replica system can be built in the event of a total loss of the original system. Backup systems should also ensure that software required to view stored information should be available as necessary.

System recovery procedures also need to be documented, to demonstrate that they 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 were 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 that 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 enhance the evidential weight of the stored information.

System audit trails should be kept of all backup activity, which should include details of any problems incurred during the procedure. These audit trails (see 4.5.3) should also detail all data file recovery activities, including a description of any problems experienced during the recovery procedures.

Backup media should be held in a secure and environmentally suitable off-site location, as part of the organization's business continuity plan (see 5.14.9). It should be transferred to the off-site location in a secure and managed manner (see 5.10.4).

Where the structure of the data files held on backup media is different from that of the originals, the structure of the backup files should be detailed in the system description manual.

Procedures for checking that data file integrity has not been compromised during a recovery from system failure should be documented.

Media used for system and data backup are not necessarily suitable for long-term storage. Media suppliers usually provide information regarding recommendations for long-term storage, including testing regimes. Alternatively, if such specific information is not available, general recommendations can often be found in national or international standards, such as BS 4783-2.

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 has been removed, is lost or is non-operational.

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

KEY ISSUES

> Back up and verify all electronic files and associated information, including metadata and audit trails, at regular intervals.

> Store backup media in a secure and environmentally suitable off-site location, with secure transfers.

> Test the viability and readability of backup media regularly, as part of a business continuity (or other) plan.

> Retention of information on backups should not be allowed to compromise the effective, timely, controlled disposal of information under the terms of the retention and disposal procedures and retention schedules.

5.14.9 Business continuity planning

From time to time, problems arise with information management systems that require emergency procedures to be implemented in order to recover from them. 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) should be implemented.

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

KEY ISSUE

> Business continuity plans should include procedures that ensure (or highlight any actual or potential issues concerning) the maintenance of the integrity of stored information, during and after an incident.

5.15 System maintenance

5.15.1 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, created by or stored within it is affected.

Preventive maintenance should be carried out regularly, in accordance with the supplier’s recommendations. These procedures may be performed by system operators or by specialized service personnel.

A maintenance log should be kept, stating the preventive and corrective maintenance procedures completed. The log should include information regarding system downtime and details of action taken.

Where system access controls can be bypassed during maintenance of hardware and/or software, personnel performing such processes should be strictly controlled, monitored and audited.

KEY ISSUE

> There should be documented procedures used for regular preventive and corrective maintenance (including appropriate security issues).

5.15.2 Scanning systems

It is of specific importance in a document scanning system that document scanners are maintained in accordance with the manufacturer's specifications, in order that image quality is maintained. This maintenance will also improve the reliability of document feed systems, thus reducing the risk of 'double feeds'.

Where document scanning is implemented, procedures described under 'Quality control' (see 5.2.3.7) 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 may be used to confirm, at any later date, that any poor quality images were not due to malfunction of the system. If there is, however, any deterioration in the output quality, the implementation of appropriate corrective maintenance procedures will be necessary.

KEY ISSUE

> Regular maintenance, particularly of scanners, will improve system quality and performance.

5.16 External service provision

5.16.1 General

DEFINITIONS

Outsource – where part of an organization's function is given to a third-party service provider to run on its behalf

NOTE: This often includes a transfer of workers to the third party.

Insource – where functions are brought into a company from a third-party service provider. This could be the reversal of a previous outsource deal or where there is a benefit from a combined operation

Co-source – where two or more organizations form a partnership, joint venture or similar alliance for mutual benefit

Many organizations have found it cost-effective to subcontract non-core parts of their business to third parties that handle these functions for many organizations. Examples of such ventures include building security, building maintenance, catering, computer application support, computer maintenance and basic business operations such as opening post and scanning documents, indexing, data conversion and similar services.

Many companies have transformed themselves into organizations that look to grow by taking on these functions. The term commonly used to define this transfer of business function is 'outsourcing'.

When considering outsourcing, the requirement for the use of, and compliance with BS 10008, should be communicated to the third party as part of the due diligence process and continued compliance with BS 10008 stated as a requirement, and thus included in service contracts.

Where the contract does not require that the third party must comply with all relevant requirements of BS 10008 and recommendations of the Code, the organization's inspection procedures on services provided should be such that no assumptions are made regarding the completeness, quality and accuracy of the services. Should this be the case, non-compliant third-party procedures should be identified. Internal procedures for the inspection of the output from the third party should be used to ensure that such non-compliances do not result in potential or actual compromise to the authenticity, integrity and/or availability of the electronic information.

The procedures and recommendations in this section 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 third-party service provider, the resulting information stored by the client will be equally admissible legally, as if the work had been done wholly within the organization;
- the organization can demonstrate compliance, many years after the event, even if the third-party 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 (see 5.10).

Where the third-party service provider uses procedures that comply with BS 10008, the organization should hold a copy of, or have access to when required, the third-party service provider's compliance documentation.

Where the organization is already operating in compliance with BS 10008, all processes to be transferred to the third-party service provider will need to be re-evaluated and the outsource company will need to demonstrate compliance. The basic processes, such as security and staff references, should already have been completed as part of the discussions leading up to the business proposal. The compliance workbook, BIP 0009, should be used to assess the complete changed environment.

Depending on the organization's audit and compliance rules, new compliance issues could arise just by using a third-party service provider. These could include the requirement for greater levels of security, encryption of data links and staff vetting procedures.

When working with third-party service providers, it is important to pay particular attention to the management interface between the parties and to the resolution processes for issues and problems arising between the parties, particularly where compliance with BS 10008 may be compromised.

Where IT-related services are being subcontracted, it may be appropriate to use the provisions contained in BS ISO/IEC 20000-1:2011, *Information technology — Service management — Part 1: Service management system requirements*. This standard specifies the best practice requirements for a set of interrelated management processes and forms the basis of an audit-managed service. Advice on implementing this standard can be found in BS ISO/IEC 20000-2:2012, *Information technology — Service management — Part 2: Guidance on the application of service management systems* , which provides advice and guidance on Part 1 and offers assistance to organizations that are planning service improvements or to be audited against BS ISO/IEC 20000-1. BSI has also published BIP 0005 and BIP 0015 (2012), *IT service management self-assessment workbook*.

It is also important to ensure that all compliance documentation is updated to reflect the changed procedures. Ideally, these revisions should be completed as part of the design process and before the operation is transferred.

KEY ISSUES

> Include compliance with BS 10008 in contracts with appropriate third-party service providers.

> Include the right to audit the third-party service provider's procedures to verify compliance with BS 10008 in contracts with appropriate third-party service providers.

5.16.2 Cloud services

An important delivery method for externally provided services is leveraging internet technologies in what is called 'the cloud'.

DEFINITION

The phrase 'in the cloud' is commonly used to refer to software, platforms, infrastructure, etc. that are sold 'as a service' and accessed by the user connected to the service provider's facilities and computing assets remotely through a network, commonly the internet.

The National Institute of Standards and Technology (NIST) www.nist.gov/itl/cloud/index.cfm publications regarding cloud computing identify five essential characteristics of cloud computing:

- on-demand self-service;
- broad network access;
- resource pooling;
- rapid elasticity;
- measured service.

This use of the term 'as a service' has led to a number of service models, the first three featuring in the NIST publications (the list will grow in line with imagination of marketing professionals):

- SaaS Software as a Service;
- IaaS Infrastructure as a Service;
- PaaS Platform as a Service;
- NaaS Network as a Service;
- DRaaS Disaster Recovery as a Service.

NIST have identified four deployment models:

- private cloud, operated solely for a single organization;
- public cloud, over a public network, such as the internet;
- community cloud, shared between a community of organizations with common concerns for security, jurisdiction, compliance, etc.;
- hybrid cloud, service provision utilizing a combination of two or more clouds (private, community or public).

NOTE: Many cloud service providers Terms of Service implicitly or explicitly acquire rights to the information on the service provider's infrastructure. This and the rights of the customer organization to this information at the termination of the business relationship, either planned or unplanned, should be considered.

5.16.3 Procedural considerations

Where the third-party service provider can demonstrate the implementation of procedures that conform to BS 10008, the services contract need only confirm this situation, and contain agreed procedures for checking continuing compliance.

Independent certification of compliance to BS 10008 or, for information security management, BS ISO/IEC 27001 may assist in meeting these requirements but it should be noted that such certification may not include any liability on behalf of the service provider.

Where the service provider does not share information regarding their processes and procedures or compliance with them, the organization may need to rely upon commitments made within the service provider's contractual terms and service level agreements.

The following list defines procedures and processes that need to be reviewed and included within the agreed procedures as appropriate:

- the ability to produce output and authenticated output to agreed acceptable quality standards;
- the ability to process a sample of input material (e.g. data files or paper documents for scanning) 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 ability to provide, in a readable form, a copy of the audit trails of the processing undertaken;
- the ability to create indexing data accurately (where indexing services are provided);
- that the proposed location for the work is acceptable and meets security criteria appropriate to requirements;
- that there is no significant increase in the risk of damage to transported media (e.g. paper documents to be scanned);
- that effective fire (and other dangers) detection and prevention systems are implemented;
- that appropriate security is maintained, including the trustworthiness of the intended operational staff;
- that, where appropriate, information sent for processing should be accessible at all relevant times.

KEY ISSUE

> Where the service provider operates in compliance with BS 10008, include in the service contract a statement making this compliance status a condition of contract, together with appropriate auditing requirements.

5.16.4 Transportation of documents

Where documents are physically moved from the organization to the third-party service provider's premises, opportunities exist for their loss or damage. Procedures need to be agreed to ensure that this risk is acceptable. All material being shipped should be adequately packed to avoid risk of damage, and be accompanied by a control document stating the identity and number of items included.

The third-party service provider 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 organization.

Records 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 third-party service provider, and of the signing by the persons handing over and receiving the material, respectively. The same process should be implemented on receipt of returned material.

KEY ISSUES

> Implement control procedures to track the movement of documents between the organization and the third-party service provider.

> Ensure that security arrangements during transportation are appropriate.

5.16.5 Overseas outsourcing

Where operations are to be outsourced overseas, the remote site location and its ability to perform satisfactory disaster recovery could impact on compliance with BS 10008. It is possible that an industry regulatory body may restrict the total percentage of an operation that can be outsourced overseas. In the event of problems in the location of the outsource organization (e.g. earthquake or war) making the third-party systems unavailable, the business may need to perform more processing locally. It should be noted that good third-party disaster recovery and contingency planning will manage this concern and aid compliance with BS 10008.

Specify in the contract the legal jurisdiction under which contract disputes are to be resolved (e.g. English law), and the location of any hearing (e.g. in London).

For example, if personal data are included in the information being processed overseas, compliance with the Data Protection Act 1998 (DPA) could be compromised by the transfer of data outside the European Economic Area (EEA) if appropriate safeguards are not in place, including:

- ensuring compliance with legislation and/or regulation, and in particular with the DPA, is not in jeopardy through transfer of functions offshore, by including relevant clauses in the service contract;[11]
- where necessary, seeking legal advice (including where data protection is concerned), and involving the Information Commissioner early in the outsource process definition.

KEY ISSUE

> Take additional care with overseas outsource contracts, when compared with UK- or EEA-based contracts.

5.16.6 Regulatory issues

Organizations should check to see whether their industry regulatory authorities have rules on what business functions can and cannot be outsourced. Concern may also be expressed if a regulatory body considers it cannot exercise its powers completely when a particular outsourcing contract is in place.

EXAMPLE

The FCA[12] would be concerned if a financial institution outsourced particular functions and the FCA considered that doubt could be expressed about the service provider's ability to employ adequate systems and controls.

[11] Where the processing of personal data is carried out in the USA, consideration should be given by the US organization to signing up to the 'safe harbor' principle (see www.export.gov/safeharbor/).

[12] www.fca.org.uk

It is thus important to involve your industry regulatory body as early in an outsourcing process as possible. Further, ensure that the regulatory body is kept informed about the proposed location of the service provider and where the work will be processed.

KEY ISSUE

> Identify any regulatory issues with overseas outsource contracts, and deal with them as appropriate, seeking legal advice where necessary.

5.16.7 Data integrity and long-term availability

Where data are stored remotely, especially overseas, adequate control procedures and processes will need to be implemented and audited. These procedures should include the ability to recover the information if the outsource arrangement is terminated for any reason. In order to remain compliant with BS 10008, migration paths for data will need to demonstrate that the content could not have been altered at any stage. The outsource partner or agent will have to demonstrate compliance in this area.

KEY ISSUES

> Ensure that the third-party service provider has satisfactory controls over documents sent to it and over the storage media used in the processing of them.

> The third-party service provider should have a clear media migration strategy.

> There should be adequate contingency and disaster recovery processes, including the verification that agreed service levels will be maintained until normal service is restored.

5.16.8 Competitors

Note that a third-party service provider may have outsourcing deals with your organization's competitors. Processes and procedures will need to be provided to ensure that 'Chinese walls' are in place and that data confidentiality and integrity cannot be unknowingly compromised by the third-party service provider.

It may also be important to confirm that the industry regulator is satisfied that there are no breaches of rules and guidelines.

KEY ISSUE

> Ensure that the third-party service provider has satisfactory controls over information confidentiality, particularly where they are also processing work for your competitors.

5.16.9 Use of trusted storage facilities

A secure means for detecting any tampering with a data file, or for verifying the contents of a data file, is to store a copy of the data file with a trusted third party. Where 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 required by BS 10008, and should be able and prepared to demonstrate, in the same manner as the owner, the effectiveness and security of its services to the satisfaction of a court of law.

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 third party. That third party will store the digital signature in secure conditions, such that it may be retrieved later.

KEY ISSUES

> Where trusted storage is used, implement processes to ensure that the third party receives true copies.

> Where trusted storage is used, review the third party's systems for compliance with BS 10008.

5.17 Information management testing

It is important that systems are tested prior to their introduction, and routinely throughout their life. Such tests should be designed to ensure that the information storage system meets the requirements of the organization in an effective and efficient manner.

To achieve this objective, tests need to be developed that test the various functions and features of the system. These tests should be consistent with the scope of the information storage system. Each test should have a defined objective, which may be to test a specific system function, or may involve a series of functions.

Test data should be carefully chosen, particularly where information about individuals is included. Such processing will need to be in compliance with the DPA. For further information about the use of personal data in system testing, see BIP 0002 (2009), *Data protection: Guidelines for the use of personal data in system testing*.

Testing should be designed in such a way that any risk of affecting live systems should be avoided. This is often achieved by the use of a separate test system, maybe managed under a 'model office' system. Such systems should be 'ring fenced', such that any output from the system that would normally be sent to organizations or individuals not involved in the testing is intercepted and retained within the testing environment.

Test results should be reviewed in line with the stated test objectives, and a conclusion reached as to whether the test results were acceptable or not.

Results of tests should be retained in line with the retention schedule.

KEY ISSUE

> System testing should be used to verify that the business requirements of the system are met.

6 Performance evaluation

6.1 Monitoring, measurement, analysis and evaluation

This section of the Code relates to Clause 9 of BS 10008, 'Performance evaluation'.

In order to be able to demonstrate the effectiveness of the management of stored electronic information over time, the system used will need to be monitored and reviewed from time to time.

Thus, audits of the system should be undertaken at planned intervals. Such audits may:

- follow a regular pattern (such as on an annual basis);
- be based on significant changes to the system;
- be as a result of a major system failure; and/or
- be 'without warning'.

6.2 Internal audit

6.2.1 Audit requirements

The essential characteristics of an audit should be borne in mind when developing an audit plan for a procedure or system. The essential features of an audit are that it:

- has a clearly defined purpose;
- is based on clearly defined and measurable criteria;
- is planned and undertaken competently;
- reaches a fair and objective conclusion;
- is documented in each of these respects.

The results of an audit will be an audit opinion. Such an opinion should not mislead. The results should include a clear explanation of the purpose of the audit, identify the criteria on which the audit was based and describe the key features of the audit approach (e.g. sources of audit evidence, the extent of reliance on internal controls, use of sampling techniques and any significant assumptions). They should also describe the auditor's qualifications for undertaking the work.

KEY ISSUE

> Audits should be defined, planned and undertaken against agreed criteria to enable a suitable audit opinion to be reached.

6.2.2 Audit planning

The initial stage for the planning of an audit is to determine the purpose for the audit. Such a purpose may be to identify any nonconformance to procedures, or may be to confirm conformance to procedures.

Once the purpose has been established, the scope of the audit should be identified and recorded. Such a scope may encompass the whole organization, a particular part of the organization or a particular process being undertaken within the organization.

It may also be appropriate to define audit criteria. Such a definition will provide a benchmark against which to assess a process, with the objective of establishing the extent to which the audited process complies with the criteria. Audit criteria take many forms, such as internal standards or procedures, specifications, codes of practice, industry sector standards, or contractual or statutory requirements.

Audit criteria may be internally or externally defined, and may be voluntarily, contractually or statutorily imposed. An audit may also aim to provide assurance that the criteria themselves adequately meet the requirements of the stakeholders in the audited process.

In practice, it is generally unnecessary to obtain a very high degree of assurance that audit criteria are met. It is typically sufficient for the audit to provide 'reasonable' assurance that the activity is free from 'material' error or nonconformance. However, this is not always the case. The evaluation and certification of systems for use in highly secure or safety critical systems is one example of a form of audit that aims to provide a high degree of assurance that audit criteria are met.

This level of assurance can only be provided on the basis of rigorous and time-consuming testing at commensurate cost.

An audit should be based on a quality plan, incorporating the relevant criteria, to provide a framework within which to work. This helps to ensure that all the activities that are necessary to meet the audit objectives take place in a logical sequence, are allocated to suitably skilled and experienced members of the audit team and are given appropriate weight in relation to their importance in forming an audit opinion. An audit plan also underpins discussions with the audited organization prior to the assignment, supports the agenda for the audit closure meeting and, together with the related audit reports and evidence, forms a permanent record of what has taken place.

KEY ISSUES

> Plan audits against an agreed purpose.

> The level of assurance obtained will depend upon good planning and adequate resource.

6.2.3 Audit procedures

Where a full system audit is undertaken, there should be procedures that review the following:

- that all applicable policies are being implemented in an appropriate manner;
- that established procedures are being followed;
- that appropriate technology has been implemented;
- that the technology is configured and maintained in accordance with requirements.

Where partial audits are undertaken, the procedures to be adopted should be such that the scope of the audit is followed.

There should be procedures for the recording of the audit results and of any appropriate analysis. Such results and analysis will lead to the audit opinion.

There should also be a procedure for the retention of evidence that an audit has taken place. It may be beneficial, or even necessary, to provide external bodies with evidence that competently planned and conducted audits have taken place.

KEY ISSUES

> Audits may be undertaken of the whole or of part of a system.

> Retain evidence of audits.

6.2.4 Selection of auditors

Numerous individuals or bodies undertake audits. Each will have particular reasons for doing so and particular objectives to be met. For example:

- internal auditors provide top management with assurances that policies and procedures are being complied with;
- external auditors are used where an internal audit function is not available, or where an external opinion is required by the organization;
- certification bodies are used to certify against external standards, such as BS EN ISO 9001 and BS ISO/IEC 27001;
- industry regulators such as the Financial Services Authority (FSA) will verify compliance with regulatory requirements;
- government departments will assess and report on compliance with legal requirements, particularly in the accounts and taxes fields;
- customers will monitor the activities of organizations with whom they trade.

Informal audits are also carried out routinely by line managers who review the procedures under their control, and assess these procedures for conformance to policy.

The important issue with the selection of auditors is that the audits are conducted in an objective manner, meet the audit requirements and produce impartial results.

KEY ISSUE

> Select the auditor with care, taking into consideration the required competency and independence.

6.3 Management review

6.3.1 General

In order to demonstrate that the system, including the related procedures, is continuing to provide the effective management of stored information, regular management reviews should be undertaken. Further, these reviews should be undertaken whenever significant changes to procedures and/or technology are being planned and/or have been implemented.

KEY ISSUE

> Management reviews determine whether the objectives of the system are being met.

6.3.2 Basis for review

Management reviews should be based on:

- general and specific feedback from system users;
- results of the various audits (see 6.2);
- records of procedural reviews;
- records of technology modifications.

6.3.3 Results

The management review should be used to assess whether compliance with BS 10008 is maintained. Where a risk is identified that compliance is or may be compromised, then a full review of compliance (see 6.3.4) should be undertaken.

KEY ISSUE

> Use the results of the management review to determine whether compliance with BS 10008 is maintained.

6.3.4 Demonstrating compliance

6.3.4.1 General

Information management systems should be audited on a regular basis to ensure that the provisions of BS 10008 and (where appropriate) the Code are being met and that the approved procedures are being adhered to. This audit should review audit trail data that are produced on a regular basis for evidence of ongoing, continuous compliance.

The compliance workbook, BIP 0009, may be used to enable a comprehensive assessment to be made of the user’s system for conformity to BS 10008, and subsequently to the Code, and to help identify which parts of the standard and Code are relevant to a system.

Compliance with BS 10008 and the Code should be claimed only if all recommendations, as stated in the workbook, have been met, or justifications for any non-applicable recommendations documented. Compliance with BS 10008 and the Code should be claimed via an authorized statement, examples of which are shown in 6.3.4.2.

The person identified in 2.3 as being responsible for maintaining compliance with the standard and the Code should review the results of each audit and document/implement a plan to address any non-compliances, which should be re-audited.

A record of compliance with BS 10008 and the Code should be maintained, as part of the audit trail. This record should include details of which recommendations are not considered relevant, and justifications for these decisions.

Where compliance with previous editions of BS 10008 and the Code has been claimed, copies of those editions should be retained as part of the compliance audit trail.

Where any change is made to the information management system, or to relevant procedures, which affects compliance with BS 10008 and the Code, a new audit of compliance should be undertaken.

Auditing may be carried out by authorized and trained in-house staff or by suitable third parties.

KEY ISSUES

> Use BIP 0009 to audit and document compliance with BS 10008 and the Code.

> Re-audit on a regular basis, and during major system changes.

6.3.4.2 Statement of compliance

6.3.4.2.1 General

Compliance with BS 10008 and the Code should be claimed using statements, which differ depending upon whether:

- the end user organization is claiming compliance;
- the system supplier is claiming that a system can be used in a compliant manner;
- a third party, acting as auditor, is confirming a compliance status.

Recommended text for use in compliance statements is given below. Alternative text may be used, but legal advice should be sought to ensure its suitability.

6.3.4.2.2 End user organizations

Individuals or organizations that conduct audits of their own information management systems may certify compliance via the following statement:

'[insert name of organization] confirms that the [insert name or other identification for the system] information management system is operated in compliance with BS 10008:2014.'

The statement should be signed by an officer of the organization, stating his or her position.

NOTE: The policy document should identify the individual or position within the organization authorized to sign statements of compliance with BS 10008.

6.3.4.2.3 System integrators and developers

Individuals or organizations that integrate/develop/supply information management systems may certify that their systems may be used in a compliant manner via the following statement:

'The [insert name or other identification for the system] information management system supplied by [insert name of integrator/developer/supplier] provides all facilities necessary for a user of this system for implementation in compliance with BS 10008:2014.'

The statement should be signed by an officer of the supplier organization, stating his or her position.

6.3.4.2.4 System auditors

Individuals or organizations that conduct audits of information management systems may certify compliance via the following statement:

'[insert name of auditing organization or individual] has assessed the [insert name or other identification for the system] information management system operated by [insert name of organization] for compliance with BS 10008:2014 and hereby certifies its compliance.'

The statement should be signed by an officer of the auditing organization, stating his or her position.

KEY ISSUE

> Claim compliance using an authorized statement.

7 Improvement

7.1 General

This section of the Code relates to Clause 10 of BS 10008, ‘Improvement’.

It is important to improve procedures and systems wherever appropriate. Such improvements may be to ensure that an identified issue is resolved without compromise to the stored information, and that the risk of a reappearance of the issue is minimized. The improvements may also relate to updated techniques and/or technology that will improve performance or reduce operational costs.

KEY ISSUE

> Ensure that procedures and systems are being maintained and improved by assessing the conclusions of audits.

7.2 Nonconformity and corrective actions

7.2.1 General

Any proposed improvement in procedures and/or technology should be assessed prior to its implementation to ensure that compliance with the information management and information security policies is not compromised.

Where major changes are implemented, an audit trail of the change management procedure should be produced and retained in line with the retention schedule. This audit should be completed as soon as possible after changes have been made.

Where migration and/or conversion procedures are actioned, the guidance in 5.8.7 and 5.8.8 should be followed.

7.2.2 Nonconformity

Actions should be undertaken to reduce the risk of nonconformities in relation to compliance with the information management and information security policies.

The audit procedures identified in 6.2.3 should be followed at regular intervals to identify any nonconformity at an early stage.

Where nonconformity is found, the cause of the nonconformity should be identified. An evaluation of the cause should then be completed, to identify the likelihood of the nonconformity reoccurring. Where the identified risk is significant, procedures and/or technology should be reviewed to identify ways of reducing this risk. Any identified actions from this review should be implemented.

The results of the review and details of the preventive actions taken should be documented and retained in accordance with the retention schedule.

KEY ISSUE

> Take action to reduce the risk of nonconformities occurring.

7.2.3 Corrective

From time to time, issues will arise that will or may result in a nonconformity occurring. There may, for example, be an actual or a suspected security breach. In these instances, corrective action should be taken to:

- assess and document any compromise to the authenticity, integrity and/or availability of the information affected;
- identify and action procedures for recovery from any compromise (maybe by a restore from backup);
- reassess the stored information once recovery procedures have been implemented;
- document any residual issues found by the reassessment;
- review the actions taken and identify (see 7.2.2) actions to be taken to prevent a reoccurrence of the issue.

KEY ISSUE

> Take corrective action to recover from nonconformities.

7.3 Continual improvement

7.3.1 General

There should be a mechanism for considering and acting on the findings of an audit. Although the auditor may recommend the general nature of any remedial action to correct problems uncovered by the audit, and may subsequently undertake further work to assess the extent to which remedial action has been successful, it is not the auditor's role to specify or impose particular solutions.

Organizations should review the results of all forms of audits (see 6.2.3) with an objective of continually improving the system. Such improvements can take many forms:

- system efficiency;
- system effectiveness;
- ease of operation;
- speed of operation;
- reduced risk of compromise to stored information;
- reduced risk of procedures not being followed.

KEY ISSUE

> Continual improvement should be an objective of the system.

7.3.2 Training

In order to be able to ensure that the procedures detailed in the procedures manual (see 4.5.2) are followed, staff need 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.

Training should be given to staff prior to them being given access to the appropriate parts of the system. Ongoing training should then be used to identify improvements within the system.

EXAMPLE

After specific training, the organization's group audit function took on the role of checking that procedures for the operation of all aspects of the information management system were being followed. Checks, including spot checks and scheduled reviews, were made at the same time as other audit checks were being made.

KEY ISSUE

> Training is needed to ensure that all staff who have access to the information management system adhere to agreed procedures.

Annex A Changes between the 2008 and 2014 versions

A.1 General

This annex contains a description of the main changes between the 2008 and 2014 versions of the Code.

In general terms, this version of BIP 0008-1 is the result of an overall review of the text and technical updates where appropriate. The scope of the Code has been extended to match the new scope of the revision of BS 10008, which includes the following additions and changes:

- recognition of the significant changes in recent years of how information is managed as an asset in organizations;
- inclusion of structured data within the scope;
- inclusion of the importance of stewardship of electronic information as an organizational activity;
- restructured to enable alignment with the ISO Management System Standards structure as defined in Annex SL of the ISO/IEC Directives, Part 1.

A.2 Editorial changes

BSI have published a document to assist with implementation of the ISO/IEC Directives, Annex SL entitled BIP 0140 (2013), *Understanding the new ISO management system requirements* which states (on page 15) 'familiar concepts such as PLAN-DO-CHECK-ACT and preventive actions have disappeared and replaced by new ones'.

However, PDCA is typically incorporated into new and revised management system standards, only now the PDCA cycle is not explicitly addressed in the standard as was the case in older versions.

There is a mapping of the PDCA cycle in the structure of management system standards that comply with Annex SL:

Section on Annex SL compliant standard	Topic	PDCA phase
Clause 4	Context of the organization	Plan
Clause 5	Leadership	Plan
Clause 6	Planning	Plan
Clause 7	Support	Plan
Clause 8	Operations	Do
Clause 9	Performance evaluation	Check
Clause 10	Improvement	Act

Table 5 – PDCA to ISO/IEC Annex SL mapping

It is therefore suggested that PDCA should not be ignored; its value is still recognized and appreciated.

The main editorial change from the 2008 edition of the Code is the restructuring to enable it to be used in conjunction with management system standards such as BS EN ISO 9001 and BS ISO/IEC 27001.

A.3 Technical changes

The main technical change from the 2008 version is the addition of data stored in structured databases and big data considerations.

The technical content of the sections on scanning systems has also been reviewed. Scanning software, and in particular facilities for extracting data from image files (Optical Character Recognition, OCR) has been updated in the light of improved recognition rates.

Also updated are the sections relating to write-once-read-many (WORM) technology. Since the first edition of the Code (published in 1996), WORM technology has progressed from optical media with a write-once capability (optical WORM) to systems that are controlled (by software/firmware) to operate in a WORM mode. Some jurisdictions have built a requirement for WORM technology to be used for the storage of information that may be required as evidence. Hence this capability remains an important part of this Code.

Annex B Development history

Version 1 – Published 1996

The idea for a code of practice was initially triggered by the concerns about the acceptability of electronic information (whether originals or copies of paper documents) expressed by manufacturers of computer-based information management systems and users of such systems from a wide range of industry sectors. The formation in 1993 of the Legal Images Initiative (LII) consortium of organizations gave direction and focus, but the greatest encouragement came from the users and managers of image-based electronic information management systems, who had a real and pressing need to understand and to resolve the issues that are the subject of the Code. Work was focused and directed at an early stage by the conclusions reached at an open meeting held in September 1993, which identified the requirement for a voluntary code of practice.

In parallel to the work of the LII, a group under the sponsorship of the Computing Suppliers Federation (CSF) and the UK Association for Information and Image Management (UKAIIM) produced a draft for consultation, which brought together the work of over 30 organizations (system suppliers and users). The draft focused upon the practical requirements of systems and operational procedures that would be required for a code of practice.

These two projects were amalgamated under the editorship of the UKAIIM Standards Committee, using the Five Principles developed by the Information and Document Management Association (IDMA) for the LII as the underlying structure[13] and practical requirements identified by the CSF/UKAIIM group as the detailed elements.

Version 2 – Published 1999

The 1996 edition[14] of the Code covered information stored on WORM optical storage systems. The second, and re-titled, edition extended the Code to cover any type of electronic storage medium, including those that are rewritable, and incorporated numerous improvements stemming from the experiences of and comments from users who had been applying the 1996 edition of the Code.

Version 3 – Published 2004

The third edition incorporated a major editorial revision of the text, and the inclusion of practical examples and case studies on implementation techniques. Some minor technical changes were made, particularly in the following sections:

- optical character recognition;
- outsourcing considerations;
- digital signatures and cryptography;
- file transfer (media transfer and electronic transmission).

This edition also recognized the publication of ISO/TR 15801, which was based on the second edition of the Code.

Version 4 – Published 2008

This edition was published in conjunction with the publication of BS 10008:2008.

This revision did not materially change the requirements for compliance with the Code. The revision was undertaken to update the structure of the Code to conform to the 'Plan-Do-Check-Act' model used by the British standard.

[13] Published as PD 0010:1998, *Principles of good practice for information management* – now withdrawn.

[14] PD 0008:1996, *Code of practice for legal admissibility of information stored on electronic document management systems*

Version 5 – Current version

This edition was published in conjunction with the publication of BS 10008:2014.

This revision did not materially change the requirements for compliance with the Code. However, the Code now has an increased scope, including the use of cloud services for service delivery, and the issues related to the management of big data.

The revision was undertaken to update the structure of the Code to conform to the requirements of Annex SL of the ISO/IEC Directives. This aligns the structure of the Code with that of the standard, which itself has a structure aligned to many of the ISO management system standards.

Annex C Definitions

For the purposes of this document, the definitions in BS ISO 12651-1:2012, *Electronic document management — Vocabulary — Part 1: Electronic document imaging*, marked (*), and those in this annex apply.

For other terms, reference can be made to the *Dictionary of Computer Science* (a compilation of the mulitple parts contained in BS ISO/IEC 2382 (series), *Information technology — Vocabulary*).

audit

systematic, independent and documented process for obtaining audit evidence and evaluating it objectively to determine the extent to which the audit criteria are fulfilled

NOTE 1: An audit can be an internal audit (first party) or an external audit (second party or third party), and it can be a combined audit (combining two or more disciplines).

NOTE 2: 'Audit evidence' and 'audit criteria' are defined in ISO 19011.

[BS ISO/IEC 27000:2014, *Information technology — Security techniques — Information security management systems — Overview and vocabulary*]

audit trail

data that allow the reconstruction of a previous activity, in its correct chronological place, or which enables the attributes of a change (such as date/time or operator) to be recorded

NOTE: The list can be generated by a computer system (for computer system transactions) or manually (usually for manual activities).

audit trail data

information stored in the audit trail (see definition)

authenticity

property that an entity is what it is claims to be

[BS ISO/IEC 27000:2014]

availability

property of being accessible and usable upon demand by an authorized entity

[BS ISO/IEC 27000:2014]

big data

collection of data sets so large and complex that it becomes difficult to process using conventional database management tools or data processing applications

NOTE: There is currently a lack of recognized international standards definition of the term 'big data'.

bitmap

digitized image where each pixel is represented by one bit, representing black or white

bit-mapped image (*)

'image derived from a bit-map'

black border removal

removal of contiguous areas of black pixels that surround the edges of document images, and which have been created by a scanning process

cloud

model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g. networks, servers, storage, applications and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction

[NIST Special Publication 800-145 entitled 'The NIST Definition of Cloud Computing' http://csrc.nist.gov/publications/nistpubs/800-145/SP800-145.pdf]

NOTE: There is a BSI document that addresses this area BIP 0117 (2010), *Cloud computing: A practical introduction to the legal issues*.

compliance

conforming to a rule; such as a specification, policy, standard or law

compression

process of reducing the size of a data file (usually an image), by the use of a computerized algorithm that encodes redundant information into a more compact form

compression ratio (*)

'ratio between the number of bits in an electronic image before and after compression'

co-source

where two or more organizations form a partnership, joint venture or similar alliance for mutual benefit

data

series of digital or analogue signals or encoded characters stored or transmitted electronically, or marks (e.g. writing, printed characters or graphics) on paper or microform, that are intended to convey information

NOTE: The essential distinction between data and information is that data do not need to have any meaning attached to them. Data become information via context; the same data file (see definition) may be considered as containing different information depending on the context: the data in the same file may be considered as containing ASCII characters as far as one computer program is concerned; another may regard the same data as a set of names and addresses; the user may know that these are customer names and addresses.

data file (*)

'related data handled as a discrete unit'

data record

item of data in a specified format as part of a data file

database

1. collection of interrelated data stored together in one or more computerized files

2. collection of data organized according to a conceptual structure describing the characteristics of the data and the relationships among their corresponding entities, supporting one or more application areas

[BS ISO/IEC 2382-1:1993]

3. collection of data describing a specific target area that is used and updated by one or more applications

[BS ISO/IEC 29881:2010, *Information technology — Systems and software engineering — FiSMA 1.1 functional size measurement method*]

[BS ISO/IEC IEEE 24765:2010, *Systems and software engineering — Vocabulary*]

database management system (DBMS)

software system designed for the definition, creation, querying, update, and administration of databases

decompression

process of reconstituting a file that has been compressed (see 'compression') back to its original form, or to a close approximation thereof (see 'lossy compression')

deletion

process of logically removing a document from a system, often by deleting an index reference

NOTE: In this case, it is (technically) possible to undelete the document (see also 'expungement').

digital/digitized image

image consisting of pixels using ranges of discrete values

digital signature

data appended to a data file that allow the recipient of the data file to authenticate the source and the integrity of the data file

NOTE 1: A digital signature can also store the date and/or time of signing.

NOTE 2: A digital signature can also be used to demonstrate that the data file could only have originated from the purported sender.

document

information stored on media

dpi

dots per inch, a measure of resolution

edge enhancement (*)

'on an electronic image, a technique for sharpening the appearance of line edges'

NOTE: In document imaging applications, this technique generally applies to black-and-white images.

electronic signature

computer data compilation of any symbol or series of symbols executed, adopted or authorized by an individual to be the legally binding equivalent of the individual's handwritten signature

NOTE: These usually contain information regarding signature biometrics. A scanned bitmap is not an electronic signature.

electronic storage

storage medium or device used by an information management system to store information

NOTE: Electronic storage specifically includes magnetic disks, magnetic tapes, storage area networks (SANs), network attached storage (NAS), hierarchical storage management (HSM) systems and optical disks of all sorts.

encryption

reversible process of converting a data file into a secret code under the control of a key

expungement (*)

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

file (in respect of computerized data)

see ‘data file’

file (in respect of microform media)

microfilm rolls or sets of microfiche or film jackets contained within a single envelope

file (in respect of paper documents)

devices for holding documents, including inter alia, envelopes, box files and ring or other types of binder

forms removal (*)

‘system (usually software) which removes a “fixed” overlay from a digitized image, leaving only the variable data’

freeze

defines a specific point in time at which no change to the contents of a specified data file are subsequently permitted to occur

NOTE: This is a fundamental concept in the Code. After a file has been ‘frozen’ any copies made of this file should contain exactly the same data.

frozen

see ‘freeze’

governance

set of principles and processes by which an organization provides direction and oversight of business related activities

[adapted from BS ISO/IEC 27000:2014]

grey scale image (*)

‘image formed of picture elements [pixels] containing grey scale information’

hierarchical storage management (HSM) (*)

'data storage technique that involves using a number of electronic storage devices, from fast access to slow access, within which data files can be moved under system control from one device to another'

hierarchical storage system (*)

'data file storage system using a number of electronic storage devices, from fast access to slow access, within which data files can be moved under system control from one device to another'

NOTE: Storage devices can range from high cost, fast access devices to low cost, slow access devices. They are used to optimize the performance and cost characteristics of the storage and retrieval system.

information

data interpreted in an application context (see also 'data')

NOTE: For example, a string of characters may be referred to generally as data; but if these characters are understood by a person or a computer program as someone's name, then the characters convey information. Information always involves the presence of data in some format, on some medium, which could be, for example, a physical document, a document image on a screen or the contents of an electronic file.

information management system

any computer or other electronic system that stores and/or processes information in digital or analogue form

insource

where functions are brought into a company from a third-party service provider. This could be the reversal of a previous outsource deal or where there is a benefit from a combined operation

integrity

property of accuracy and completeness

[BS ISO/IEC 27000:2014]

intelligent character recognition (ICR)

technique for OCR where some degree of 'intelligence' is used in the conversion software, to determine, on the basis of other characteristics in the 'bitmapped' image, the actual character being read

key escrow service

function provided by an organization to keep secure copies of encryption keys, in a manner that requires consent from the key owner before the key is disclosed

lossless compression (*)

'data file compression technique where the decompressed image is identical to the original uncompressed image

NOTE: No information is lost during the compression and decompression process.'

lossy compression (*)

'data file compression technique where the decompressed image may not be identical to the original uncompressed image

NOTE 1: Information may be lost during the compression process.'

NOTE 2: With image files, high compression ratios (e.g. around 50:1) can be obtained with little observable image degradation.

master data management (MDM)

policies, processes, tools and governance to consistently define and manage significant data of an organization and to provide a single point of reference for that data

metadata

data about data

NOTE: For example, the context and relationships of data with other data that is necessary to render that data more understandable.

multi-functional drive system (*)

‘optical disk drive which can use both write-once read-many and rewritable optical media’

network attached storage (NAS)

server that is dedicated to nothing more than file sharing

NOTE: NAS does not provide any of the activities that a server in a server-centric system typically provides, such as email, authentication or file management.

NoSQL

database providing a mechanism for storage and retrieval of data that may be modelled in means other than the tabular relationships used in relational databases

NOTE 1: The term NoSQL is short for ’Not only SQL’ in that they may support SQL-like query languages in addition to other methods.

NOTE 2: Many NoSQL databases lack full ACID transaction support.

objective

result to be achieved

NOTE 1: An objective can be strategic, tactical or operational.

NOTE 2: Objectives can relate to different disciplines (such as financial, health and safety, and environmental goals) and can apply at different levels (such as strategic, organization-wide, project, product and process.)

NOTE 3: An objective can be expressed in other ways, for example as an intended outcome, a purpose, an operational criterion, as an information security objective or by the use of other words with similar meaning (e.g. aim, goal or target).

NOTE 4: In the context of information security management systems, information security objectives are set by the organization, consistent with the information security policy, to achieve specific results.

[BS ISO/IEC 27000:2014]

optical character recognition (OCR)

technique for the recognition of characters from a digital image

optical disk (*)

‘disk that will accept and retain information in the form of marks in a recording layer that can be read with an optical beam’

optical mark recognition (OMR)

technique for electronic detection of marks in specified locations on a digital image

organization

person or group of people that has its own functions with responsibilities, authorities and relationships to achieve its objectives

NOTE: The concept of organization includes but is not limited to sole-trader, company, corporation, firm, enterprise, authority, partnership, charity or institution, or part or combination thereof, whether incorporated or not, public or private.

[BS ISO/IEC 27000:2014]

original document

document from which a copy is made or from which an image is captured

outsource

make an arrangement where an external organization performs part of an organization's function or process

NOTE 1: This often includes a transfer of workers to the third party.

NOTE 2: An external organization is outside the scope of the management system, although the outsourced function or process is within the scope.

[BS ISO/IEC 27000:2014]

overlay

see 'forms removal'

page

single image entity, such as one side of a sheet of paper, a drawing or plan, map, photograph, transparency; or a microform 'frame'

pixel

smallest two-dimensional element of a digital image that can independently be assigned attributes such as colour and intensity

NOTE: The word is derived from 'picture element'.

record (noun)

'information created, received and maintained as evidence and information by an organization or person, in pursuance of legal obligations or in the transaction of business'

[BS ISO 15489-1:2001]

NOTE: This may be data stored in electronic form, in a data file (see definition) or on a paper or microform document.

records management

'field of management responsible for the efficient and systematic control of the creation, receipt, maintenance, use and disposition of records, including processes for capturing and maintaining evidence of and information about business activities and transactions in the form of records'

[BS ISO 15489-1:2001]

resolution (*)

'ability of a scanner or image generation device to reproduce the details of an image'

scanning

operation that converts the image of a document into a digital form, by detecting the amount of light reflected from elements of a document

schema

logical plan showing the relationships between metadata elements

NOTE: The relationships are normally through establishing rules for the use and management of metadata specifically as regards the semantics, the syntax and the optionality (obligation level) of values.

[ISO 23081-1:2006, *Information and documentation — Records management processes — Metadata for records — Part 1: Principles*]

skew

poor document alignment (rotation) during scanning

speckle

random, or quasi-random, black marks (speckles) on an image, either generated during the scanning process, or present on the original document

stewardship

responsibility for information assets within the organization

NOTE: Definition based on ISO 20121:2012, *Event sustainability management systems — Requirements with guidance for use*.

storage area network (SAN)

sub-network (typically high speed) of shared storage devices

NOTE: A storage device is a machine that contains nothing but a disk or disks for storing data and the necessary control mechanisms for storing that data. A SAN's architecture works in a way that makes all storage devices available to all servers on a local area network (LAN) or a wide area network (WAN).

system

in the Code, this always means information management system (see 'information management system'), unless specifically noted

system files (*)

'files held in a computer for use in the control and operation of the system'

Trusted Time

time stamping of information performed by, or with the direct involvement of, a party underwriting the accuracy and consistency of that time stamp

NOTE: Trusted Time stamps are used to enhance the integrity or authenticity of information to which they relate. They are often used in conjunction with cryptographic digital signatures.

vital records

records that are fundamental to the functioning of the organization

workflow

‘automation of a business process, in whole or in part, during which documents, information or tasks are passed from one participant to another for action, according to a set of procedural rules’

NOTE: This definition is taken from the Workflow Management Coalition’s ‘Terminology and Glossary’ document, WFMC-TC-1011, with permission.

write-once-read-many (WORM)

form of electronic storage medium that permits data to be stored once, read many times, but not altered

Annex D Example information management policy statement

This annex contains an example ‘information management policy statement’. It can be used as a draft upon which an organization’s policy can be based.

XYZABC Limited

ABC project

Policy document for compliance with the requirements of BS 10008:2014 Specification: Evidential weight and legal admissibility of electronic information.

Approved by:	
Name:	
Position:	
Date:	

1. Scope

This document covers the information management policies implemented within the ABC project, supported by the HIJ electronic information storage system. This policy conforms to the requirements of BS 10008:2014 Specification: Evidential weight and legal admissibility of electronic information. The HIJ electronic information storage system is described in a system description manual (Ref: SD01). Procedures for the use of the system are described in a procedures manual (Ref: PM01).

2. Information covered

Stored information covered by this policy document relates to that used in relation to all aspects of the ABC project. Documents included within the scope of this policy are detailed in the information retention policy (see Appendix A).

XYZABC Limited does not operate an information classification system, as all information is regarded as having the same security level.

3. Storage media and file formats

All stored information as detailed above is held in a format and on media as described in Appendix A.

4. Standards

All electronic information within XYZABC Limited is stored in conformance to BS 10008:2014, together with any referenced national and/or international standards.

5. Document retention

5.1 Electronic documents

This section deals with electronic documents stored on the HIJ electronic information storage system. In the case where an original document has been retained, its destruction policy is also included within this section.

The information retention policy is defined in Appendix A of this policy document. The retention periods given in Appendix A have been agreed within XYZABC Limited to cover legal and operational requirements consistent with the status of being a public limited company.

5.2 Paper documents

This section deals with original documents that are received in paper form. All original documents are securely and confidentially destroyed, in accordance with procedures documented in the procedures manual – after internal quality control procedures have been successfully completed, with the following exceptions:

1. original documents in the following categories are kept for the relevant retention period:
   - poor quality documents for which a satisfactory image has not been obtained (see PM01/Image quality), or which required significant enhancement;
   - documents with physical amendments that have not been captured;
   - documents where fraud is suspected;
2. original documents not owned by XYZABC Limited are returned to the originator;
3. original documents should be retained where necessary for legal reasons (e.g. contracts).

6. Destruction policy

The procedure for the destruction of paper and electronic documents is detailed in the procedures manual. All original documents covered by this policy document are destroyed using these procedures.

No paper original documents are destroyed until the electronic version has been quality checked.

Electronic documents are only routinely destroyed during media migration procedures. A certificate of destruction is produced to record the reason for the destruction of the original document, to meet statutory and regulatory requirements.

Paper or electronic documents are not destroyed where litigation is pending or in progress. Where data protection legislation is applicable, auditable procedures for access to and destruction of personal information subject to the requirements of the legislation will be documented.

7. Responsibilities

This policy document should be reviewed annually under the control of the Company Secretary. Where changes are agreed, they are to be implemented using the change control procedures (Ref: CC01).

This policy, and any revisions, should be approved by the Board of Directors of XYZABC Limited prior to its implementation.

The maintenance of compliance with BS 10008 is the responsibility of the Head of Internal Audit.

8. Legal advice sought

XYZABC Limited has sought and obtained agreement for the information retention policy detailed in Appendix A of this document.

9. Duty of care

XYZABC Limited has a duty to keep secure and accurate original documentation, or authentic copies of them. This is achieved by:

- implementing this policy document;
- implementing an information security policy (Ref: ISP01);
- ensuring that only trained staff have access to the system;
- ensuring that acceptable quality control procedures are implemented;
- ensuring that XYZABC Limited's legal advisers are consulted, and appropriate actions taken;
- ensuring that appropriate audit trails are created and retained.

Appendix A Information retention policy

This policy relates to XYZABC Limited's HIJ electronic information storage system only.

The retention and destruction policies are in accordance with XYZABC Limited's archiving strategy (Ref: ARC01).

This policy is reviewed annually, as part of the XYZABC Limited's risk assessment programme.

Document type	Media type*	Format	Retention period**	Responsibility
Financial reports	Rewritable optical disk	TIFF	7 years	Company Secretary
Supplier documentation	CD/Paper***	PDF	Longer of: 1. life of equipment + 10 years, or 2. longest related information storage	Archivist
Internally produced documents	Magnetic tape	ASCII text	End of project + 10 years	Originator
Letters, incoming and outgoing	1. Paper 2. WORM	1. Paper 2. TIFF	1. 1 month after post scanning quality control 2. End of project + 10 years	Archivist
Purchase orders	3. Paper 4. WORM	3. Paper 4. TIFF	3. 1 month after post-scanning quality control 4. End of project + 10 years	Archivist
Procedures	WORM	PDF	Longest related	Operations
Audit reports	WORM	PDF	Longest related	Audit

* Paper originals that are archived to WORM and/or CD media are routinely destroyed (under the control of the scanning department) after successful completion of quality control procedures (see 5.2).

** Documents on CD and WORM media are only destroyed during a media migration process; magnetic and paper records are destroyed annually as appropriate.

*** These documents both paper (two copies) and CD form. The retention period relates to the CD version only, which is the controlled version.

Appendix B Legal advice sought

Advice has been sought from:

- XYZABC Limited’s legal department;
- external auditors;
- external legal advisers;
- industry regulators;
- HM Revenue & Customs.

Annex E Records management

BS ISO 15489 is the British Standard on records management. It deals with the issues surrounding the standardization of records management policies and procedures, to ensure that appropriate attention is given to all records stored by an organization. The following table lists the general contents of BS ISO 15489-1:2001, and maps them across to the Code.

BS ISO 15489-1:2001		BIP 0008-1 (2014)	
4	Benefits of records management	–	Introduction
5	Regulatory environment	Annex H	Legal issues
6	Policy and responsibilities	1	Information management planning
7.1	Principles of records management programmes	1	Information management planning
7.2	Characteristics of a record	–	General
8	Design and implementation of a records system	2	Implementing and operating
9	Records management processes and controls	2	Implementing and operating
10	Monitoring and auditing	3	Monitoring and reviewing
11	Training	4	Maintaining and improving

Table 6 – Mapping BS ISO 15489-1 to the Code

Annex F Application of controls

This annex contains details of specific applications that can be managed under the control of this Code of Practice, and indicates which sections of the Code are relevant and which may not be relevant.

Use Cases

In-House Scanning

- Scanning of information from physical documents addressed by compliance scope within the enterprise

Outsourced Scanning

- Scanning of information from physical documents addressed by compliance scope as a service provided by a third-party organization outside the enterprise.

Service Provider – Scanning

- Third party providing scanning service to an organization. Whilst these service providers may have their own policies (etc.), the information management policies (etc.) are those of the information owning organization not the service provider.

Structured Databases

- Information addressed by compliance scope within the enterprise held in structured database(s); this information itself may be structured or unstructured.

Big Data User

- Information addressed by compliance scope within the enterprise regarded as 'big data'; this information itself may be structured or unstructured.

In-House Storage

- Storage of information addressed by compliance scope within the enterprise.

Outsourced Storage

- Storage of information addressed by compliance scope as a service provided by a third-party organization outside the enterprise.

Service Provider – Storage

- Third party providing storage services to an organization; using cloud or other communications methods; storage may also be of physical media.
  Whilst these service providers may have their own policies (etc.), the information management policies (etc.) are those of the information owning organization not the service provider.

Table 7 - Applicability Matrix

BS 10008:2014 Section	BS 10008:2014 Section Title	In-House Scanning	Outsourced Scanning	Service Provider–Scanning	Structured Databases	Big Data	In-House Storage	Outsourced Storage	Service Provider – Storage
4	Context of the organization	✓	✓	✗	✓	✓	✓	✓	✗
4.1	General	✓	✓	✗	✓	✓	✓	✓	✗
4.2	Issues	✓	✓	✗	✓	✓	✓	✓	✗
4.3	Requirements	✓	✓	✗	✓	✓	✓	✓	✗
4.4	Boundaries and applicability	✓	✓	✗	✓	✓	✓	✓	✗
5	Leadership	✓	✓	✗	✓	✓	✓	✓	✗
5.1	Leadership and commitment	✓	✓	✗	✓	✓	✓	✓	✗
5.2	Policy statements	✓	✓	✗	✓	✓	✓	✓	✗
5.2.1	General	✓	✓	✗	✓	✓	✓	✓	✗
5.2.2	Electronic storage policy statement	✓	✓	✗	✓	✓	✓	✓	✗
5.2.3	Electronic transfer policy statement	✓	✓	✗	✓	✓	✓	✓	✗
5.2.4	Information security policy	✓	✓	✗	✓	✓	✓	✓	✗
5.3	Roles and responsibilities of workers	✓	✓	✗	✓	✓	✓	✓	✗
5.4	Legal and regulatory environment	✓	✓	✓	✓	✓	✓	✓	✓
6	Planning	✓	✓	✓	✓	✓	✓	✓	✓
6.1	Actions to address risks and opportunities	✓	✓	✓	✓	✓	✓	✓	✓
6.1.1	General	✓	✓	✓	✓	✓	✓	✓	✓
6.1.2	Risk assessment	✓	✓	✓	✓	✓	✓	✓	✓
6.1.3	Risk treatment	✓	✓	✓	✓	✓	✓	✓	✓
6.2	Objectives and achievements	✓	✓	✓	✓	✓	✓	✓	✓
7	Support	✓	✓	✓	✓	✓	✓	✓	✓
7.1	Resources	✓	✓	✓	✓	✓	✓	✓	✓
7.2	Competence	✓	✓	✓	✓	✓	✓	✓	✓
7.3	Awareness	✓	✓	✓	✓	✓	✓	✓	✓

BS 10008:2014 Section	BS 10008:2014 Section Title	In-House Scanning	Outsourced Scanning	Service Provider–Scanning	Structured Databases	Big Data	In-House Storage	Outsourced Storage	Service Provider – Storage
7.4	Reporting and communication	✓	✓	✓	✓	✓	✓	✓	✓
7.5	Documented information	✓	✓	✓	✓	✓	✓	✓	✓
7.5.1	General	✓	✓	✓	✓	✓	✓	✓	✓
7.5.2	Procedural documentation	✓	✓	✓	✓	✓	✓	✓	✓
7.5.3	Audit trails	✓	✓	✓	✓	✓	✓	✓	✓
8	Operation	✓	✓	✓	✓	✓	✓	✓	✓
8.1	Information capture	✓	✓	✓	✓	✓	✗	✗	✗
8.1.1	General	✓	✓	✓	✓	✓	✗	✗	✗
8.1.2	Importing	✗	✗	✗	✓	✓	✗	✗	✗
8.1.3	Document scanning	✓	✓	✓	✗	✗	✗	✗	✗
8.1.4	Data extraction	✓	✓	✓	✓	✓	✗	✗	✗
8.1.5	Metadata capture	✓	✓	✓	✓	✓	✓	✓	✓
8.2	Self-modifying files	✗	✗	✗	✓	✓	✓	✓	✗
8.3	Compound documents	✗	✗	✗	✓	✓	✓	✓	✗
8.4	Information in structured databases	✗	✗	✗	✓	✓	✓	✓	✓
8.4.1	General	✗	✗	✗	✓	✓	✓	✓	✓
8.4.2	Big Data considerations	✗	✗	✗	✓	✓	✓	✓	✓
8.5	Version control	✓	✓	✓	✓	✓	✓	✓	✓
8.6	Storage systems	✓	✓	✓	✓	✓	✓	✓	✓
8.6.1	Storage technology	✓	✓	✓	✓	✓	✓	✓	✓
8.6.2	Migration	✗	✗	✗	✓	✓	✓	✓	✓
8.6.3	Storage file formats	✓	✓	✓	✓	✓	✓	✓	✗
8.6.4	Conversion	✓	✓	✓	✓	✓	✓	✓	✗
8.6.5	Compression	✓	✓	✓	✓	✓	✓	✓	✓
8.7	Information transfer	✗	✓	✓	✓	✓	✓	✓	✓

BS 10008:2014 Section	BS 10008:2014 Section Title	In-House Scanning	Outsourced Scanning	Service Provider–Scanning	Structured Databases	Big Data	In-House Storage	Outsourced Storage	Service Provider – Storage
8.7.1	General	✗	✓	✓	✓	✓	✓	✓	✓
8.7.2	Transmission	✗	✓	✓	✓	✓	✓	✓	✓
8.7.3	Message transmission systems	✗	✗	✗	✓	✓	✗	✓	✗
8.8	Indexing and other metadata	✓	✓	✓	✓	✓	✗	✗	✓
8.9	Authenticated output procedures	✓	✓	✗	✓	✓	✓	✓	✗
8.10	Identity	✓	✓	✓	✓	✓	✓	✓	✓
8.11	Information retention and disposition	✓	✓	✗	✓	✓	✓	✓	✗
8.11.1	Retention	✓	✓	✗	✓	✓	✓	✓	✗
8.11.2	Disposition	✓	✓	✗	✓	✓	✓	✓	✗
8.12	Information security procedures	✓	✓	✓	✓	✓	✓	✓	✓
8.12.1	General	✓	✓	✓	✓	✓	✓	✓	✓
8.12.2	Access rights	✓	✓	✓	✓	✓	✓	✓	✓
8.12.3	Encryption	✓	✓	✓	✓	✓	✓	✓	✓
8.12.4	Digital signatures	✓	✓	✓	✓	✓	✓	✓	✓
8.12.5	Back-up and recovery	✓	✓	✓	✓	✓	✓	✓	✓
8.12.6	Business continuity planning	✓	✓	✓	✓	✓	✓	✓	✓
8.13	System maintenance	✓	✓	✓	✓	✓	✓	✓	✓
8.14	External service provision	✗	✓	✗	✓	✓	✗	✓	✗
8.14.1	Procedures	✗	✓	✗	✓	✓	✗	✓	✗
8.14.2	Compliance	✗	✓	✗	✓	✓	✗	✓	✗
8.14.3	Security in transfer	✗	✓	✗	✓	✓	✗	✓	✗
8.14.4	Overseas	✗	✓	✗	✓	✓	✗	✓	✗
8.15	Information management testing	✗	✓	✗	✓	✓	✗	✓	✗
9	Performance evaluation	✓	✓	✓	✓	✓	✓	✓	✓

BS 10008:2014 Section	BS 10008:2014 Section Title	In-House Scanning	Outsourced Scanning	Service Provider–Scanning	Structured Databases	Big Data	In-House Storage	Outsourced Storage	Service Provider – Storage
9.1	Monitoring, measurement, analysis and evaluation	✓	✓	✓	✓	✓	✓	✓	✓
9.2	Internal audit	✓	✓	✓	✓	✓	✓	✓	✓
9.3	Management review	✓	✓	✓	✓	✓	✓	✓	✓
10	Improvement	✓	✓	✓	✓	✓	✓	✓	✓
10.1	Nonconformity and corrective actions	✓	✓	✓	✓	✓	✓	✓	✓
10.2	Continual improvement	✓	✓	✓	✓	✓	✓	✓	✓

Annex G References

BSI publications

British Standards Institution, London. BSI Publications are available from Customer Services, Sales Department, 389 Chiswick High Road, London W4 4AL. Tel: 020 8996 9001; Fax: 020 8996 7001

PD 0018:2001, *Information management systems — Building systems fit for audit*

Secure/Multipurpose Internet Mail Extensions (S/MIME) Version 3.1 Message Specification. Available at: www.ietf.org/rfc/rfc3851.txt

Standards

BS ISO/IEC 2382 (series), *Information technology — Vocabulary*

BS 4783 (series), *Storage, transportation and maintenance of media for use in data processing and information storage*

BS 7083:1996, *Guide to the accommodation and operating environment for information technology (IT) equipment*

BS 10008:2014, *Evidential weight and legal admissibility of electronic information — Specification*

BS 10012:2009, *Data protection — Specification for a personal information management system*

BS EN ISO 9000, *Quality management systems*

BS EN ISO 9001:2008, *Quality management systems — Requirements*

BS EN ISO 14001:2004, *Environmental management systems — Requirements with guidance for use*

BS EN ISO 19011:2011, *Guidelines for auditing management systems*

BS ISO 12651-1:2012, *Electronic document management — Vocabulary — Part 1: Electronic document imaging*

BS ISO 12653 (series), *Electronic imaging — Test target for the black-and-white scanning of office documents*

BS ISO 15489-1:2001, *Information and documentation — Records management*

BS ISO 19005-1:2005, *Document management — Electronic document file format for long-term preservation — Part 1: Use of PDF 1.4 (PDF/A- 1)*

BS ISO/IEC 20000-1:2011, *Information technology — Service management — Part 1: Service management system requirements*

BS ISO/IEC 20000-2:2012, *Information technology — Service management — Part 2: Guidance on the application of service management systems*

BS ISO 20121:2012, *Event sustainability management systems — Requirements with guidance for use*

BS ISO 23081-1:2006, *Information and documentation — Records management processes — Metadata for records — Part 1: Principles*

BS ISO/IEC 27000:2014, *Information technology — Security techniques — Information security management systems — Overview and vocabulary*

BS ISO/IEC 27001:2013, *Information technology — Security techniques — Information security management systems — Requirements*

BS ISO/IEC 27002:2013, *Information technology — Security techniques — Code of practice for information security controls*

BS ISO/IEC 27005:2011, *Information technology — Security techniques — Information security risk management*

BS ISO/IEC 29881:2010, *Information technology — Systems and software engineering — FiSMA 1.1 functional size measurement method*

BS ISO 31000:2009, *Risk management — Principles and guidelines*

BS ISO 32000-1:2008, *Document management — Portable document format — Part 1: PDF 1.7*

BS ISO/IEC IEEE 24765:2010, *Systems and software engineering — Vocabulary*

The following standard is available from ISO:

ISO/TR 15801:2009, *Document management — Information stored electronically — Recommendations for trustworthiness and reliability*

Guidance documents

BIP 0002 (2009), *Data protection: Guidelines for the use of personal data in system testing*

BIP 0005 (2011), *A manager's guide to service management*

BIP 0008-2 (2014), *Evidential weight and legal admissibility of information transferred electronically — Code of practice for the implementation of BS 10008*

BIP 0008-3 (2014), *Evidential weight and legal admissibility of linking electronic identity to information — Code of practice for the implementation of BS 10008*

BIP 0009 (2014), *Evidential weight and legal admissibility of electronic information — Compliance workbook for use with BS 10008*

BIP 0015 (2012), *IT service management self-assessment workbook*

BIP 0071 (2013), *Guidelines on requirements and preparation for ISMS certification based on ISO/IEC 27001*

BIP 0072 (2013), *Are you ready for an ISMS audit based on ISO/IEC 27001?*

BIP 0073 (2013), *Guide to the implementation and auditing of ISMS controls based on ISO/IEC 27001*

BIP 0074 (2006), *Measuring the effectiveness of your ISMS implementations based on ISO/IEC 27001*

BIP 0140 (2013), *Understanding the new ISO management system requirements*

BIP 0117 (2010), *Cloud computing: A practical introduction to the legal issues*

ISO/IEC Directives, Part 1 — *Consolidated ISO Supplement — Procedures specific to ISO, Fifth edition 2014*

Other publications

Auld, Lord Justice (2001) *A Review of the Criminal Courts of England and Wales*. Available at: www.criminal-courts-review.org.uk/

ISO (1997) *Dictionary of Computer Science: The Standardized Vocabulary*

Anti-terrorism, Crime and Security Act 2001, London: HMSO

Business Names Act 1985, London: HMSO

Civil Evidence Act 1995, London: HMSO

Civil Evidence Act (Northern Ireland) 1971, London: HMSO

Civil Evidence (Scotland) Act 1988, London: HMSO

Companies Act 1985 and 1989, London: The Stationery Office

Computer Misuse Act 1990, London: HMSO

Consumer Protection Act 1987, London: HMSO

Copyright, Designs and Patents Act 1988, London: HMSO

Criminal Justice Act 1988 and 1991, London: HMSO

Criminal Justice and Police Act 2001, London: HMSO

Criminal Justice and Public Order Act 1994, London: HMSO

Criminal Justice (Evidence, Etc.) (Northern Ireland) Order 1988, London: HMSO

Criminal Procedure and Investigations Act 1996, London: HMSO

Data Protection Act 1998, London: HMSO

Electronic Communications Act 2000, London: HMSO

Evidence Act (Northern Ireland) 1939, London: HMSO

Financial Services and Markets Act 2000, London: HMSO

Freedom of Information Act 2000, London: HMSO

Human Rights Act 1998, London: HMSO

Insolvency Act 2000, London: HMSO

Latent Damage Act 1986, London: HMSO

Limitation Act 1980, London: HMSO

Obscene Publications Act 1959 and 1964, London: HMSO

Police and Criminal Evidence Act 1984, London: HMSO

The Police and Criminal Evidence (Northern Ireland) Order 1989, London: HMSO

Protection from Harassment Act 1997, London: HMSO

Protection of Children Act 1978, Protection of Children

Race Relations Act 1976, London: HMSO

Regulation of Investigatory Powers Act 2000, London: HMSO

Sex Discrimination Act 1975, London: HMSO

Sex Offenders Act 1997 London: HMSO

Statute Law Revision Act (Northern Ireland) 1953, London: HMSO

Taxes Management Act 1970, London: HMSO

The Telecommunications (Lawful Business Practice)(Interception of Communications) Regulations 2000, London: HMSO

Value Added Tax Act 1994, London: HMSO

Lord Chancellor's Code of Practice on the Management of Records, Issued under section 46 of the Freedom of Information Act 2000, November 2002. Available at: www.justice.gov.uk/downloads/information-access-rights/foi/foi-section-46-code-of-practice.pdf

Lord Chancellor's Department, Law Officers Department (2002) *Justice for All*. Cm 5563, London: The Stationery Office

Hawley Committee (1995) *Information as an Asset: The Board Agenda; a Consultative Report*, London: KPMG/Impact Group

Rochester Institute of Technology, Process Ink Gamut Chart. Available from Rochester Institute of Technology, T & E Center, One Lomb Memorial Drive, Rochester, New York 14623, USA

Workflow Management Coalition (WfMC) (1999) *Workflow Management Coalition Terminology & Glossary*. WFMC–TC–1011, Winchester: WfMC

Financial Reporting Council, *UK Corporate Governance Code (2014)* (formerly The Combined Code). Available at: www.frc.org.uk/Our-Work/Codes-Standards/Corporate-governance/UK-Corporate-Governance-Code.aspx

Financial Reporting Council, *UK Stewardship Code (2012)*. Available at: www.frc.org.uk/Our-Work/Codes-Standards/Corporate-governance/UK-Stewardship-Code.aspx

Annex H Legal issues

H.1 Background

The Code describes means by which it may be demonstrated at any time, in a manner acceptable to a court of law, that the contents of a specific data file created or existing within a computer system have not changed since the time of storage (the defined time), and that where such a data file contains a digitized image of a physical original document, the digitized image is a true facsimile of that original document. The issue being addressed is essentially one of authenticity and integrity.

This defined time fixes the time at which the contents of the data file were frozen for the purposes of storage and future retrieval. For example:

- the time at which a scanned and digitized image of an original document was stored on a computer system;
- the time when a word-processed document was stored.

Other versions of the information may legitimately develop, for example revision of a contract, or update of a spreadsheet or word-processed document, and these can similarly be frozen at defined times. In these cases the new versions are treated as new data files.

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

The Code 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, a physical original document or contained voice and/or video information.

Irrespective of issues of legal admissibility or evidential weight, the Code defines best practice for electronic storage of business or other information. As such, complying with its recommendations is of value to organizations even when evidential issues are not relevant.

KEY ISSUE

> The Code defines best practice for electronic storage, such that authenticity and integrity can be easily demonstrated, in a manner appropriate in a court of law.

H.2 Information management and legal admissibility

There has been much discussion about the value of information stored electronically when required as evidence in a court of law, or for other purposes. It is crucial that a discipline is commonly agreed so that the value of this information as evidence can be maximized. The Code has been developed to address this need.

The Code is for use as a basic reference document. It covers information stored on any type of electronic storage medium, including hierarchical storage systems that allow for migration between storage media either automatically or under operator control. It also covers information management systems where information is transmitted via local or wide area networks. Thus, electronic messages, such as those produced by EDI and email systems, can be stored under the controls of the Code. In applications where information is stored as data records within a file, then the Code may also be applied at the data record level.

Compliance with the Code does not guarantee legal admissibility. It defines best practice.

The Code covers issues such as system planning, implementation, initial loading and procedures for the use of the system, including workflow. It pays particular attention to setting up authorized procedures and subsequently being able to demonstrate, in a court of law, that these procedures have been followed.

Procedures are defined that need to be implemented in order to comply with the Code. However, it does not follow that information stored on systems that do not comply with the Code is not or will not be legally admissible.

KEY ISSUE

> Compliance with the recommendations of the Code will enable the organization to demonstrate that best practice is being followed.

H.3 Weight of evidence and document destruction

In order to maximize evidential weight, it is important to determine, in advance, how information would be presented to a court of law, and whether weight of evidence or courtroom tactics could be unduly influenced by the destruction of the original document, by the information management system or by the access control systems.

One of the major reasons for using electronic information storage is the potential ability to destroy original paper documents. Implementing a destruction policy, however, is not without risk. The electronic document may be rejected in court, and the original document requested. This risk will be reduced where an organization can demonstrate that it has complied with its documented procedures and processes, and that these are aligned with the requirements of the Code.

The Code cannot give a definitive recommendation regarding the destruction of original documents. It is for the organization to assess the risk, and make an appropriate business decision. The organization's legal adviser will be able to give an opinion on which types of document are most likely to be disputed in court. There may be different considerations for civil (on the balance of probabilities) and criminal (beyond reasonable doubt) law.

KEY ISSUE

> Compliance with the recommendations of the Code increases the weight of evidence, and potentially enables original document destruction where appropriate.

H.4 Authenticity

It is important to be able to demonstrate that the computer system has been functioning properly (i.e. according to agreed procedures) in order to authenticate data stored on the system.

Arguments over admissibility of information as evidence can lead to an investigation into the system from which the information came, the method of storage, operation and access control, and even into the computer programs and source code. It may be necessary to satisfy the court that the information is stored in a 'proper' manner. This could be a tactic used to try to discredit the evidence and to make inadmissible, or reduce the evidential weight of, that evidence and any similarly stored information that is produced. Questionable hardware reliability, for example, could be used to discredit the information management system. This could call the whole system into question and cause information stored within it to be ruled inadmissible.

The implementation of documented procedures for storage, maintenance and monitoring access to the information will minimize this risk.

KEY ISSUE

> Compliance with the recommendations of the Code enables authenticity to be demonstrated.

H.5 Originals and copies

Data may be moved around, within and between computer systems. Whether or not the data is modified in this process, it may be necessary to be able to demonstrate that the data are ‘original’, in the sense that they have not been changed since they entered or were created within a system, or were ‘frozen’ for the purpose of storage. In the case of scanned documents, the ‘original’ is the original document itself. In the case of a document created by a word processor, the original may be considered as the file first created by the author. However, whether a data file or a physical document is an original may depend on the context. For example, a letter received from a second party may be considered as an original for the purpose of proving receipt, whilst a copy of that letter sent to a third party may be considered as an original for the purpose of proving what the third party received. The same is true for electronic data, but with the difference that a copy contains exactly the same information as the original data file. Without information being available regarding the time at which a file was first stored, it can be impossible to distinguish between the original data file and a copy.

The key recommendation is to be able to demonstrate the sequence of events, so that a distinction between a copy and the original can always be made. If a data file is copied, so that two or more identical data files subsequently exist, the method used to determine which file is the ‘original’ and which is the ‘copy’ may be questioned. If an original file, or a copy of it, is maliciously altered, the perpetrator may claim that the amended file is the ‘original’. Therefore, it may be necessary to be able to determine the date and time at which a file was first stored in order to determine whether the submitted data file is actually the ‘original’ or the ‘copy’. The Code recommends procedures for identifying a copy of an original.

In the context of a system where an original data file may be moved from one medium to another, the distinction between the original and a copy can be a fine one. When a data file is moved from one medium to another, the original on the first medium no longer exists as far as the system is concerned, because the space on the first medium, previously occupied by the data file, may be reused. The data file as it now exists on the second medium is effectively the equivalent of the original, and may need to be considered as such as far as the Code is concerned. The Code recommends procedures for authenticating the data file, however often it may be moved.

When considering original documents that are to be scanned and stored in e-form, the issue of copy versus original can be more difficult to deal with, and a pragmatic approach is necessary. If an original document contains a signature that is handwritten in ink on ‘original’ letterhead paper, then organizations may be correct in assuming that it is an original. If an original document has been produced in multiple copies on a laser printer, for example, the minutes of a meeting, each such document is in principle an ‘original’. The organization needs to implement a policy for dealing with original documents, and the need to determine authenticity, prior to their being stored on the information management system.

KEY ISSUE

> Where only a copy exists, it will be treated as ‘best evidence’. There may be a need to demonstrate the authenticity of copies stored.

H.6 Born digital environment

Increasingly, information used in business is held in a ‘dynamic, digital’ form. Typical of this type of system are websites where electronic trading processes are controlled. Information stored on such systems may vary under system control or by operator intervention. In such cases, details will need to be stored about the information content of the system at particular instances (e.g. when each purchase is made) in the system life. There may also be a requirement to retain specific metadata (see 5.7.2) to augment the evidential weight of the information in question.

KEY ISSUE

> Where dynamic data are involved, a method of capturing 'snapshots' of the information at particular times will be needed

H.7 Digitized images

Because of the great importance of digitized images of original documents in many applications, the Code provides detailed guidance on document image capture and the committal to storage of these images, to ensure that 'true facsimiles' of original documents are created and stored. Original documents may exist on paper or on microform. Particularly, the Code describes procedures:

- to ensure that all necessary information is captured as accurately as possible from the external, original document;
- to demonstrate that the captured image has not been changed since its creation, or, where change is permitted by the application, the precise nature of such change (e.g. conversion from colour to grey scale, or to black and white, de-skewing of the original image, or cropping of the original image to removed unwanted parts);
- to capture contextual information about the original document (metadata) in order to reinforce the evidential value of the captured electronic version of the original document.

NOTE: While indexing is required for virtually all data files, contextual information is of particular importance for digitized images because the image itself cannot be searched for relevant text, as can other types of data file.

KEY ISSUE

> Demonstrating the authenticity of electronic images (scanned documents) requires appropriate and auditable scanning procedures and processes.

H.8 Photocopies, microfilm and electronic images

Electronic images of original paper documents will be treated as secondary evidence in the same manner as photocopies or microform images. There is also the possibility of a reduction in evidential value whenever a copy is made. As an example, if a signature is disputed, then the original document, upon which the original ink is preserved, may hold more weight in a court of law than the copy, because a forensic expert may derive more evidence from the paper and ink used. The organization needs to take these factors into account when assessing a document's suitability for scanning and subsequent destruction.

KEY ISSUE

> Electronic images will be treated as secondary evidence, just like photocopies or microform images.

H.9 Information storage

It is the responsibility of the executives of the organization to be able to produce information (either in the form of an original or a copy) when required, no matter how the organization stores this information. The authorized officer (in a company this is the company secretary or equivalent) and the manager of the information management system are responsible for this information retrieval process, and not the vendor of the system. Thus, considering the advice of the authorized officer, or equivalent, is essential before implementing any information management system, particularly when original documents are subsequently destroyed.

The procedures by which information is stored and accessed are vital in satisfying a court of law about the authenticity of a 'copy' of information and the inability to tamper with it. All copies of documents (photocopy, microform or electronic) will be treated by a court of law as secondary evidence, with a potential reduction of weight of evidence if the authenticity of the copy is questioned. For example, where the content of a document is under question, the original or a copy should be treated with equal weight, but if a signature is being disputed, then the original document is likely to carry more weight than a copy of it.

KEY ISSUE

> Information stored electronically should be managed in an appropriate manner, in accordance with the Code, to demonstrate compliance with legal and/or regulatory requirements.

H.10 Storage and access procedures

Because of the duration of storage of data, the person who 'certified' a system, or data stored on it, may not be able to give evidence in person. It is, therefore, essential that a system for monitoring and certifying is implemented to demonstrate that the integrity of the system has been maintained from the time that the data was stored.

Regular audits of the system are advised, with certificates of compliance obtained. This is in line with current procedures for microfilmed documents. Although formal affidavits will not usually be necessary, advice may need to be sought from the organization's legal adviser, particularly if original documents are to be destroyed.

It may help demonstrate the 'normal' functioning of a system if a copy of the result of the audit is stored on the system.

Where rewritable media is used, additional procedures (discussed in the Code) will need to be implemented to restrict and record access to the storage media and devices used to access it. These procedures are, in the main, common practice for most information management systems. Computer operating systems (or add-in functions) often provide suitable levels of access control, whereas application software may require modification in order to provide necessary functionality. A system audit trail, which records access to and use of storage media, will be usable as additional evidence, and thus will need to be kept for at least the same period as the information to which it relates.

Some data files, particularly those generated by word processors, may contain executable code (often referred to as 'macros'), which can have the effect of modifying a file each time it is retrieved, viewed or printed out, for example, by inserting the current date. The existence of a macro within a file means that the file cannot be frozen in the sense required by the Code. The Code recommends that such self-modifying files are not used in applications where the information may be of potential evidential value.

KEY ISSUE

> Competent auditing carried out at agreed intervals, and including documented reporting, is necessary to ensure that compliance with the Code is ongoing.

H.11 Retrieval/viewing software

In practice, in a court of law, an electronic data file will be retrieved and viewed or printed via suitable software. Regardless of the procedures followed in managing the file within the system, it is possible that different retrieval/viewing software could result in different presentations of the data, so that it might appear that the files were different. The Code makes specific recommendations for dealing with this issue.

KEY ISSUE

> The ability to retrieve and view stored information in an authenticated manner is key to demonstrating authenticity.

H.12 United Kingdom legislation

H.12.1 Civil Evidence Act 1995

In England and Wales, the Civil Evidence Act 1995 is of major importance to the Code. This Act resolved many of the problems of legal admissibility per se that had arisen over evidence generated by or held on computers. It shifted the argument away from admissibility to the evidential value or weight of a document.

The previous complicated rules on computer evidence have been abolished and there is now no special provision for computer evidence. It will also be easier to demonstrate the authenticity of documents that are admissible in evidence, by producing the original or a copy, regardless of how many removes there are between the original and the copy. The court will still need to be satisfied as to the authenticity of the document, and the organization should therefore put in place procedures to demonstrate this. As can be seen from the Act, section 9(1) uses the words ‘may be received’. From this statement, it is evident that the court will have the power to exclude documents in which trust has been compromised.

Sections 8 and 9 of the Act are as follows:

8 **Proof of statements contained in documents**

(1) Where a statement contained in a document is admissible as evidence in civil proceedings, it may be proved:
  (a) by the production of that document, or
  (b) whether or not that document is still in existence, by the production of a copy of that document or of the material part of it, authenticated in such a manner as the court may approve.

(2) It is immaterial for this purpose how many removes there are between a copy and the original.

9 **Proof of records of business or public authority**

(1) A document which is shown to form part of the records of a business or public authority may be received in evidence in civil proceedings without further proof.

(2) A document shall be taken to form part of the records of a business or public authority if there is produced to the court a certificate to that effect signed by an officer of the business or authority to which the records belong.

Much civil evidence legislation throughout the world makes similar provisions.

H.12.2 Police and Criminal Evidence Act 1984

The essential provisions of the Civil Evidence Act 1995 are also included in the proposed replacement to the Police and Criminal Evidence Act 1984 (PACE). These are expected to replace the current cumbersome requirements of section 69 of the Act. Section 69 relates to computer evidence.

The revised PACE codes of practice were published in 2008. There are sections in these codes (in particular in Code B[15]) that relate to electronic evidence.

EXAMPLE

Code B, section 7.6

'If an officer considers information stored in any electronic form and accessible from the premises could be used in evidence, they may require the information to be produced in a form:

- that can be taken away and in which it is visible and legible; or
- from which it can readily be produced in a visible and legible form'.

In his Criminal Court Review (2001) Lord Justice Auld recommended 'that the law should, in general, move away from technical rules of inadmissibility to trusting judicial and lay fact finders to give relevant evidence the weight it deserves'.

This has been accepted and is incorporated in the 2002 White Paper Justice for All – A White Paper on the Criminal Justice System. Sections 4.52 and 4.53 of the White Paper include:

'4.52 The current rules of evidence, which determine what evidence the court can take into account, are difficult to understand and complex to apply in practice. There has been growing public concern that evidence relevant to the search for truth is being wrongly excluded.

4.53 Magistrates, judges and juries should be trusted to give appropriate evidence the weight it deserves when they exercise their judgement…'

That it is relevant to include this in the Code is emphasized by Section 4.61 of the same White Paper, which affirms '…the right approach is that…where records have been properly compiled by businesses, then the evidence should automatically go in, rather than its admissibility being judged…'. This is followed by the observation that 'This is close to the approach developed in civil proceedings.'

H.12.3 Data Protection Act 1998

It may be necessary to amend, delete or expunge specific information from information management systems, owing to a court order and/or to meet the requirements of the Data Protection Act 1998.

[15] http://police.homeoffice.gov.uk/operational-policing/powers-pace-codes/pace-code-intro/

The information management system may need to have facilities to delete or destroy information. It should also have the facility to amend incorrect data, or remove irrelevant data, held in contravention of the Data Protection Act 1998.

For further information on the Data Protection Act 1998, and how it can be implemented, see BS 10012:2009, *Data protection — Specification for a personal information management system.*

H.12.4 Freedom of Information Act 2000

At the time of publication of the Code, the majority of UK public authorities were required to publish an approved publication scheme, and make this information available to the public on request. As from January 2005, all information held by the public body will need to be available to the public on request, unless it falls within a specified list of exemptions.

In order to comply with this 'all information' access requirement, a code of practice was published by the Lord Chancellor's Office,[16] detailing the need to implement effective records management practices. The Code states that authorities 'should seek to conform to the provisions of BSI DISC PD0008 – A Code of Practice for Legal Admissibility and Evidential Weight of Information Stored Electronically (2nd edition) – especially for those records likely to be required as evidence'. (PD 0008 has now been superseded by BIP 0008.)

H.12.5 List of pertinent UK legislation

This list has been compiled from a number of sources, the main ones being the previous version of the Code, and The National Archives, among other sources.

It should be noted that new legislation is continually being introduced and, therefore, this list cannot be exhaustive. Therefore, it is essential that the organization using the Code continually monitors the legal scene and reacts to changes in legislation appropriately.

Consideration should also be given to European Directives that are not yet covered by legislation in the relevant part of the UK, but can nevertheless be enforced. Compliance with the Code will improve the evidential weight of information used in any dispute, whatever the legislative position.

- Anti-terrorism, Crime and Security Act 2001;
- Business Names Act 1985;
- Civil Evidence Act 1995;
- Civil Evidence Act (Northern Ireland) 1971;
- Civil Evidence (Scotland) Act 1988;
- Companies Acts 1985 and 1989;
- Computer Misuse Act 1990;
- Consumer Protection Act 1987;
- Copyright, Designs and Patents Act 1988;
- Criminal Justice Act 1988 and 1991;
- Criminal Justice and Police Act 2001;
- Criminal Justice and Public Order Act 1994;
- Criminal Justice (Evidence, Etc.) (Northern Ireland) Order 1988;
- Criminal Procedure and Investigations Act 1996;
- Data Protection Act 1998;
- Evidence Act (Northern Ireland) 1939;
- Financial Services and Markets Act 2000;
- Freedom of Information Act 2000;
- Human Rights Act 1998;
- Insolvency Act 2000;
- Latent Damage Act 1986;

[16] www.justice.gov.uk/downloads/information-access-rights/foi/foi-section-46-code-of-practice.pdf

- Limitation Act 1980;
- Obscene Publications Act 1959 and 1964;
- Police and Criminal Evidence Act 1984;
- The Police and Criminal Evidence (Northern Ireland) Order 1989;
- Protection from Harassment Act 1997;
- Protection of Children Act 1978;
- Race Relations Act 1976;
- Regulation of Investigatory Powers Act 2000;
- Sex Discrimination Act 1975;
- Sex Offenders Act 1997;
- Statute Law Revision Act (Northern Ireland) 1953;
- Taxes Management Act 1970;
- The Telecommunications (Lawful Business Practice) (Interception of Communications) Regulations 2000;
- Value Added Tax Act 1994.

Evidential weight and legal admissibility of information stored electronically

Code of practice for the implementation of BS 10008

Evidential weight and legal admissibility of information stored electronically – Code of practice for the implementation of BS 10008 is primarily concerned with the authenticity, integrity and availability of electronically stored information, to the demonstrable levels of certainty required by an organization. It is particularly applicable where this stored information may be used as evidence in disputes inside and outside the legal system.

Now in its fifth edition, the book provides a valuable framework and guidelines that identify key areas of good practice for the implementation and operation of such electronic storage systems.

This fifth edition is technically similar to the fourth edition, with an extension of its scope to include information stored in databases and other electronic systems. It has also been restructured in recognition of the publication of BS 10008:2014, and can be considered to be a guide to the implementation of the standard in relation to information stored electronically.

This publication is the first part of BIP 0008. The other two parts are:
BIP 0008-2 (2014), *Evidential weight and legal admissibility of information transferred electronically — Code of practice for the implementation of BS 10008;*
BIP 0008-3 (2014), *Evidential weight and legal admissibility of linking electronic identity to information — Code of practice for the implementation of BS 10008.*

This book provides guidance on how your organization can:

- improve reliability of, and confidence in, stored information;
- maximize the evidential weight which a court or other body may assign to presented information;
- provide confidence in inter-company trading;
- provide confidence to external inspectors (for example, regulators and auditors) that the organization's information and business practices are robust and reliable.

Peter Howes

Peter Howes is Director and Principal Consultant for Group 5 Training Limited and is a specialist in the practical issues of legal and regulatory compliance, governance, electronic communications and information security, with over 40 years' relevant experience in the business application of information systems. For the last 20 years, Peter has worked with BSI to develop the full range of evidential weight and legal admissibility publications and has delivered a wide range of workshops on evidential weight as well as email records management, email and the law, information security and the law with BSI.

Alan Shipman

Alan Shipman is Managing Director and Principal Consultant for Group 5 Training Limited. He has been involved in Document Imaging Standards for over 20 years, specializing in user aspects. Alan is Chairman of the BSI Document Imaging Applications committee, convenor of the ISO Document Imaging Quality sub-committee and a member of the UKAIIM Standards Committee. Alan has presented BSI Training Workshops on the practical implementation of BIP 0008, as well as speaking on the subject at events in the information management, archives and records management fields and also at industry specific events in educational, engineering, health care, financial, legal, local government and system supplier fields.

bsi.

BSI Group Headquarters
389 Chiswick High Road
London W4 4AL
www.bsigroup.com

The British Standards Institution is incorporated by Royal Charter

BSI order ref: BIP 0008/1

9 780580 856761