

| | | | |
|--|---|--|--|
| COMMITTEE DRAFT | | Reference number: | |
| ISO/IEC 1 st CD 27050-1 | | ISO/IEC JTC 1/SC 27 N14101 | |
| Date: 2014-07-28 | | Supersedes document SC 27 N13304 | |
| THIS DOCUMENT IS STILL UNDER STUDY AND SUBJECT TO CHANGE. IT SHOULD NOT BE USED FOR REFERENCE PURPOSES. | | | |
| ISO/IEC JTC 1/SC 27 Information technology - Security techniques Secretariat: Germany (DIN) | | Circulated to P- and O-members, and to technical committees and organizations in liaison for comments by: 2014-10-28 Please submit your comments via the online balloting application by the due date indicated. | |
| ISO/IEC 1 st CD 27050-1 | | | |
| Title: Information technology -- Security techniques – Electronic discovery — Part 1: Overview and concepts | | | |
| Project: 1.27.107.01 (27050-1) | | | |
| Explanatory Report | | | |
| Status | SC 27 Decision | Reference documents | |
| | | Input | Output |
| WG 4 Study Period on Electronic discovery | 12th WG 4 meeting, May 2012, resolutions 6, 16, 19 (N11309). | | Call f. contr. (N11306). |
| NWIP on Electronic discovery Preliminary draft | 13th WG 4 meeting, Oct. 2012, resolutions 27, 29 (N11941). | SoContr. (N11627); Report (N11668). | Report (N11954); Text f. NWIP (N11955); Text f. preliminary draft (N11956). |
| ISO/IEC NP 27050 1 st WD | 14th WG 4 meeting, April 2013, resolutions 18, 25, 34 (N12740). | SoV (N12633); ISO/TC 46/SC 11 liaison com. (N12222). | DoC (N12678); Liaisons to: CDFS (N12636); CSA (N12299); EuroCloud (N12639); ETSI ISG ISI (N12638); FIRST (N12640); ISO/TC 46/SC 11 (N12646); DoC (N12678); Text f. 1 st WD (N12679). |
| ISO/IEC NP 27050-1 1 st WD | 15th WG 4 meeting, October 2013, resolutions 15, 20, 26 (N13271). | SoCom (N12679). | Liaisons to: CDFS (N13275); ISO/TC 46/SC 11 (N13278). Justification (N13326); Request/endorsement f./on subdivision (N13543 / N13886); DoC (N13303); Text f. 1 st WD (N13304). |
| ISO/IEC 27050-1 1 st CD | 16th WG 4 meeting, April 2014, resolutions 19, 23 (in N12740); 26 th SC 27 Plenary, April 2014, Resolutions 1, 16, Del. of Auth. f. 1 st DIS Resolution 8 (N14200). | SoCom (N13749); CSA 11 liaison com. (N13802). | Liaisons to: CDFS (N14041); CSA (N14247); Interpol (N14045); ISO/TC 46/SC 11 (N14052); ITU-T SG 17 (N14245); DoC (N14098); Text f. 1 st CD (N14099). |
| CD Registration and Consideration | | | |
| In accordance with resolution 1 (contained in SC 27 N14200) of the 26 th SC 27 Plenary meeting held in Hong Kong, China, 14 th – 15 th April 2014 the attached document has been registered with the ISO Central Secretariat as a 1 st Committee Draft (CD) and is hereby circulated for a 3-month 1 st CD letter ballot closing by | | | |
| 2014-10-28 | | | |
| MEDIUM: http://isotc.iso.org/livelink/livelink/open/jtc1sc27 | | | |
| NO. OF PAGES: 1 + 34 | | | |

Information technology — Security techniques — Electronic discovery — Part 1: Overview and concepts

Warning

This document is not an ISO International Standard. It is distributed for review and comment. It is subject to change without notice and may not be referred to as an International Standard.

Recipients of this draft are invited to submit, with their comments, notification of any relevant patent rights of which they are aware and to provide supporting documentation.

Copyright notice

This ISO document is a working draft or committee draft and is copyright-protected by ISO. While the reproduction of working drafts or committee drafts in any form for use by participants in the ISO standards development process is permitted without prior permission from ISO, neither this document nor any extract from it may be reproduced, stored or transmitted in any form for any other purpose without prior written permission from ISO.

Requests for permission to reproduce this document for the purpose of selling it should be addressed as shown below or to ISO's member body in the country of the requester:

Secretariat of ISO/IEC JTC 1/SC 27
DIN German Institute for Standardization
DE-10787 Berlin

Tel. + 49 30 2601 2652

Fax + 49 30 2601 1723

E-mail krystyna.passia@din.de

Web <http://www.jtc1sc27.din.de/en> (public web site)

<http://isotc.iso.org/isotcportal/index.html> (SC 27 documents)

Reproduction for sales purposes may be subject to royalty payments or a licensing agreement.

Violators may be prosecuted.

Contents

Page

| | |
|--|----|
| Foreword | v |
| 0 Introduction..... | vi |
| 1 Scope | 1 |
| 2 Normative references | 1 |
| 3 Terms and definitions | 1 |
| 4 Symbols and abbreviated terms | 6 |
| 5 Overall ISO/IEC 27050 structure and overview | 7 |
| 5.1 Purpose and structure | 7 |
| 5.2 Overview of Part 1: Overview and concepts | 7 |
| 5.3 Overview of Part 2: Guidance for governance and management of electronic discovery | 7 |
| 5.4 Overview of Part 3: Code of practice for electronic discovery | 7 |
| 5.5 Overview of Part 4: ICT readiness for electronic discovery | 8 |
| 6 Overview of electronic discovery | 8 |
| 6.1 Background..... | 8 |
| 6.2 Basic concepts | 8 |
| 6.3 Objectives of electronic discovery | 9 |
| 6.4 Electronic discovery foundation for success | 9 |
| 6.4.1 General | 9 |
| 6.4.2 Competency | 10 |
| 6.4.3 Candour | 10 |
| 6.4.4 Cooperation | 10 |
| 6.4.5 Completeness | 10 |
| 6.4.6 Proportionality | 10 |
| 6.5 Planning and budgeting an electronic discovery project | 10 |
| 7 Electronically Stored Information (ESI)..... | 11 |
| 7.1 General | 11 |
| 7.2 Common types of ESI | 12 |
| 7.2.1 General | 12 |
| 7.2.2 Active data | 12 |
| 7.2.3 Inactive data | 12 |
| 7.2.4 Residual data | 12 |
| 7.2.5 Legacy data | 13 |
| 7.3 Common sources of ESI | 13 |
| 7.3.1 General | 13 |
| 7.3.2 Custodian data sources | 13 |
| 7.3.3 Non-custodian data sources | 14 |
| 7.3.4 Potentially excluded sources of ESI | 15 |
| 7.4 ESI representations..... | 16 |
| 7.4.1 Native formats..... | 16 |
| 7.4.2 Near native formats | 16 |
| 7.4.3 Image (near paper) formats | 16 |
| 7.4.4 Paper..... | 16 |
| 7.5 Non-ESI as part of discovery | 17 |
| 8 Electronic discovery process | 17 |

| | | |
|--------------|---|----|
| 8.1 | Overview | 17 |
| 8.2 | Identification phase | 18 |
| 8.3 | Preservation phase..... | 19 |
| 8.4 | Collection phase | 19 |
| 8.5 | Processing phase | 20 |
| 8.6 | Review phase | 20 |
| 8.7 | Analysis phase..... | 21 |
| 8.8 | Production phase..... | 21 |
| 9 | Additional considerations..... | 21 |
| 9.1 | Information management..... | 21 |
| 9.2 | ESI presentation..... | 22 |
| 9.3 | Chain of custody and provenance | 22 |
| 9.4 | Protection of ESI | 22 |
| 9.4.1 | General..... | 22 |
| 9.4.2 | Long-term retention of ESI | 22 |
| 9.4.3 | Maintaining ESI confidentiality..... | 23 |
| 9.4.4 | Disposition of ESI | 23 |
| Annex A | (informative) Potentially relevant standards | 24 |
| A.1 | General..... | 24 |
| A.2 | Issues with investigative concepts and terminology..... | 26 |
| Bibliography | | 28 |

Foreword

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

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

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

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

ISO/IEC 27050-1 was prepared by Technical Committee ISO/TC JTC 1, *Information technology*, Subcommittee SC 27, *Security techniques*.

ISO/IEC 27050 consists of the following parts, under the general title *Information technology — Security techniques — Electronic discovery*:

- *Part 1: Overview and concepts*
- *Part 2: Guidance for governance and management of electronic discovery*
- *Part 3: Code of practice for electronic discovery*
- *Part 4: ICT readiness for electronic discovery*

0 Introduction

This International Standard provides an overview of electronic discovery and describes related terminology, concepts, and processes, which are intended to be leveraged by the other ISO/IEC 27050 parts. It is relevant to both non-technical and technical personnel involved in some or all of the electronic discovery activities. It is important to note that this guidance is not intended to contradict or supersede local jurisdictional laws and regulations.

Electronic discovery often serves as a driver for investigations (covered in ISO/IEC 27041, ISO/IEC 27042, and ISO/IEC 27043) as well as evidence acquisition and handling activities (covered in ISO/IEC 27037). In addition, the sensitivity and criticality of the data sometime necessitate protections like storage security to guard against data breaches (covered in ISO/IEC 27040).

It should be noted that this International Standard is not a reference or normative document for regulatory and legislative security requirements. Although it emphasizes the importance of these influences, it cannot state them specifically, since they are dependent on the country, the type of business, etc.

Information technology — Security techniques — Electronic discovery — Part 1: Overview and concepts

1 Scope

Electronic discovery is the process of discovering pertinent Electronically Stored Information (ESI) or data by one or more parties involved in an investigation and any resulting actions. This International Standard provides an overview of electronic discovery. In addition, it defines related definitions and describes the concepts, including, but not limited to identification, preservation, collection, processing, review, analysis, and production of ESI. This International Standard also identifies other relevant standards (e.g. ISO/IEC 27037) and how they relate to, and interact with, electronic discovery activities.

This International Standard is relevant to both non-technical and technical personnel involved in some or all of the electronic discovery activities, and it is not intended to contradict or supersede local jurisdictional laws and regulations, so care should be exercised to ensure compliance with the prevailing jurisdictional requirements.

2 Normative references

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

ISO/IEC 27000:2014, *Information technology — Security techniques — Information security management systems — Overview and vocabulary*

ISO/IEC 27001:2013, *Information technology — Security techniques — Information security management systems — Requirements*

ISO/IEC 27002:2013, *Information technology — Security techniques — Code of practice for information security controls*

3 Terms and definitions

For the purposes of this document, the terms and definitions given in ISO/IEC 27000, ISO/IEC 27005, and the following apply.

3.1 analysis

element of an *electronic discovery* (3.10) process focused on evaluating *Electronically Stored Information* (3.11) for content and context to identify facts, relationships, key patterns, and other features that can lead to improved understanding of an ESI (3.11) corpus

Note 1 to entry: Content and context can include key patterns, topics, people and discussions.

3.2

chain of custody

possession, movement, handling, and location of material from the time it is obtained to the time it is presented in a matter

3.3

collection

element of an *electronic discovery* (3.10) process focused on gathering *Electronically Stored Information* (3.11) and other related material

3.4

custodian

person or organization that has ownership, custody or administrative control of *Electronically Stored Information* (3.11)

3.5

data breach

compromise of security that leads to the accidental or unlawful destruction, loss, alteration, unauthorized disclosure of, or access to protected data transmitted, *stored* (3.29) or otherwise processed

[SOURCE: ISO/IEC 27040, 3.7.]

3.6

data integrity

property that data has not been altered or destroyed in an unauthorized manner

[SOURCE: ISO 7498-2:1989, 3.3.21.]

3.7

digital evidence

information or data, *stored* (3.29) or transmitted in binary form, that may be relied on as evidence

[SOURCE: ISO/IEC 27037:2012, 3.5.]

3.8

discovery

process by which each party gains information held by an adverse party concerning a matter

Note 1 to entry: *Discovery* is also the disclosure of facts, documents, *Electronically Stored Information* (3.11) and tangible objects by an adverse party.

3.9

electronic archive

long-term repository of *Electronically Stored Information* (3.11)

Note 1 to entry: *Electronic archives* can be on-line, and therefore accessible, or off-line and not easily accessible.

Note 2 to entry: Backup systems (e.g., tape, virtual tape, etc.) are not intended to be *electronic archives*, but rather data protection systems (i.e., recovery mechanisms for disaster recovery and business continuity).

3.10

electronic discovery

process that includes the *identification* (3.12), *preservation* (3.19), *collection* (3.3), *processing* (3.20), *review* (3.24), *analysis* (3.1), and *production* (3.21) of *Electronically Stored Information* (3.11)

Note 1 to entry: Although *electronic discovery* is often considered a legal process, its use is not limited to the legal domain.

3.11

Electronically Stored Information

data or information of any kind and from any source, whose temporal existence is evidenced by being *stored* (3.29) in or on any electronic medium

Note 1 to entry: *ESI* includes traditional e-mail, memos, letters, spreadsheets, databases, office documents, presentations and other electronic formats commonly found on a computer. *ESI* also includes system, application and file-associated *metadata* (3.14) such as timestamps, revision history, file type, etc.

Note 2 to entry: Electronic medium can take the form of, but is not limited to, storage devices and storage elements.

[SOURCE: ISO/IEC 27040, 3.16.]

3.12

identification

element of an *electronic discovery* (3.10) process focused on locating potential sources and the criteria for selecting potentially relevant *Electronically Stored Information* (3.11)

3.13

legal hold

process of suspending the normal disposition or processing of records and *Electronically Stored Information* (3.11) as a result of current or anticipated litigation, audit, government investigation or other such matters

Note 1 to entry: The issued communication that implements the legal hold may also be called a "hold," "preservation order," "suspension order," "freeze notice," "hold order," or "hold notice."

3.14

metadata

data that defines and describes other data

[SOURCE: ISO/IEC 11179-1:2004, 3.2.16.]

3.15

native file

electronic document in a *native format* (3.16)

Note 1 to entry: Native files are frequently proprietary.

3.16

native format

organization and representation of data and *metadata* (3.14) that an operating system or application uses when data is *stored* (3.29)

Note 1 to entry: *Native formats* typically contain the most complete representation of the data. While it is often possible to convert this data to other formats, there can be a loss of information (e.g., *metadata* is stripped) or modification of the information.

Note 2 to entry: In many circumstances or jurisdictions *native format* is that format in which a document is *stored* (3.29) or used in the normal course of its use/business.

3.17

non-volatile storage

storage (3.28) that retains its contents even after power is removed

[SOURCE: ISO/IEC 27040, 3.30.]

3.18
potential digital evidence

digital evidence (3.7) that has not yet been admitted as such in a court of law or other legal proceeding

[SOURCE: ISO/IEC 27043, 3.12.]

3.19
preservation

element of an *electronic discovery* (3.10) process focused on ensuring that *Electronically Stored Information* (3.11) is protected against inappropriate alteration or destruction

Note 1 to entry: In some matters or jurisdictions, there can be requirements to prevent *spoliation* (3.27) of Electronically Stored Information (3.11).

3.20
processing

element of an *electronic discovery* (3.10) process focused on extracting *Electronically Stored Information* (3.11) and converting it, if necessary, to forms more suitable for *review* (3.24) and *analysis* (3.1)

3.21
production

element of an *electronic discovery* (3.10) process focused on delivering or making available *Electronically Stored Information* (3.11)

Note 1 to entry: *Production* can also include getting *Electronically Stored Information* (3.11) in appropriate forms and using appropriate delivery mechanisms.

Note 2 to entry: *Production* can be to any person or organization.

3.22
production file format

organization and representation of data and *metadata* (3.14) that is presented to a requesting party

3.23
provenance

information that documents the origin or source of *Electronically Stored Information* (3.11), any changes that may have taken place since it was originated, and who has had custody of it since it was originated

3.24
review

element of an *electronic discovery* (3.10) process focused on evaluating *Electronically Stored Information* (3.11) for relevance and privilege

Note 1 to entry: In some matters or jurisdictions, *Electronically Stored Information* (3.11) that is considered privileged can be excluded from *production* (3.21).

3.25
sanitize

process to remove information from media such that data recovery is not possible at a given level of effort

[SOURCE: ISO/IEC 27040, 3.38.]

Note 1 to entry: Clear, purge, and destruct are actions that can be taken to *sanitize* storage media.

3.26**search**

use of various methods for identifying and finding potentially relevant *Electronically Stored Information* (3.11) based on certain criteria

Note 1 to entry: The actual process of *searching* can take many forms (e.g., keyword, fuzzy, Boolean, phonic, synonym, etc. searches).

Note 2 to entry: The content considered a match for a particular *search* may not be an exact match to the criteria.

3.27**spoliation**

act of making or allowing change(s) to the *potential digital evidence* (3.18) that diminishes its evidential value

[SOURCE: ISO/IEC 27037:2012, 3.19.]

3.28**storage**

device, function or service supporting data entry and retrieval

[SOURCE: ISO/IEC 27040, 3.43.]

3.29**stored**

process that results in data being recorded on *volatile storage* (3.30) or *non-volatile storage* (3.17)

[SOURCE: ISO/IEC 27040, 3.50.]

3.30**volatile storage**

storage (3.28) that fails to retain its contents after power is removed

[SOURCE: ISO/IEC 27040, 3.53.]

1 **4 Symbols and abbreviated terms**

| | | |
|----|--------|---|
| 2 | CD | Compact Disc |
| 3 | DVD | Digital Video Disc |
| 4 | EDMS | Electronic Document Management System |
| 5 | ERMS | Electronic Records Management System |
| 6 | ESI | Electronically Stored Information |
| 7 | FTP | File Transfer Protocol |
| 8 | HR | Human Resources |
| 9 | HTML | HyperText Markup Language |
| 10 | IM | Instant Messaging |
| 11 | ICT | Information and Communications Technology |
| 12 | IT | Information Technology |
| 13 | NAS | Network Attached Storage |
| 14 | OCR | Optical Character Recognition |
| 15 | PII | Personally Identifiable Information |
| 16 | RAM | Random Access Memory |
| 17 | TAR | Technology Assisted Review |
| 18 | WebDAV | Web Distributed Authoring and Versioning |
| 19 | XML | Extensible Markup Language |

20

5 Overall ISO/IEC 27050 structure and overview

5.1 Purpose and structure

ISO/IEC 27050 is a multi-part International Standard that provides requirements and guidance for the process of discovering pertinent Electronically Stored Information (ESI) or data by one or both parties involved in an investigation and any resulting actions. Figure 2 provides notional architecture of this multi-part International Standard.

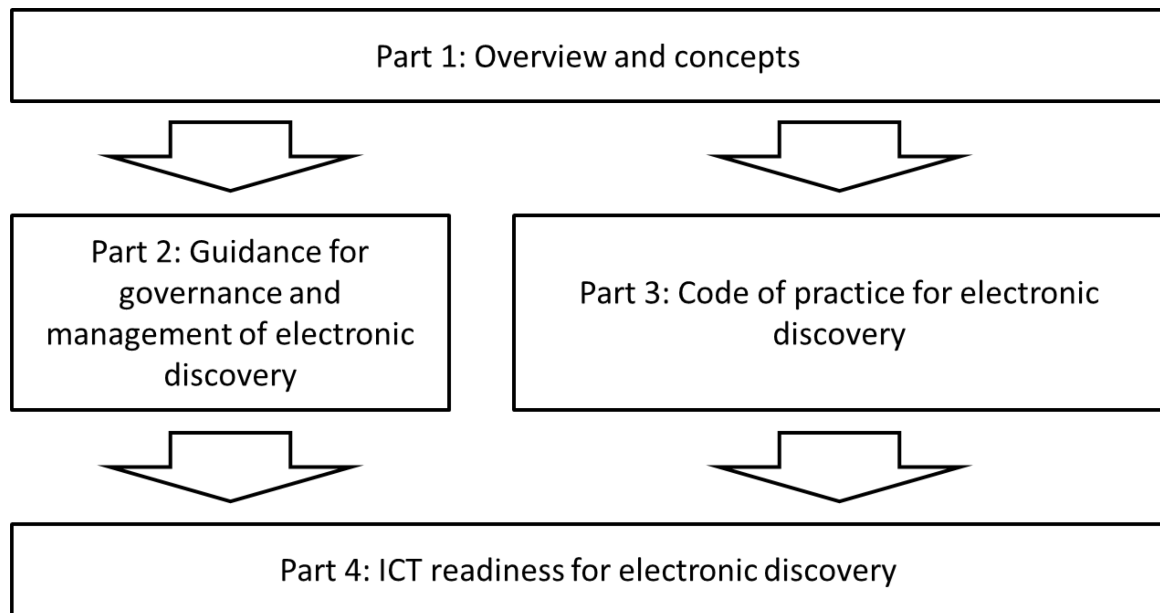


Figure 1 — ISO/IEC 27050 architecture

5.2 Overview of Part 1: Overview and concepts

Part 1 (this document) provides an overview of electronic discovery, introducing relevant terminology, concepts, and processes. Part 1 is an informative document.

5.3 Overview of Part 2: Guidance for governance and management of electronic discovery

Part 2 addresses how personnel at senior levels within an organization can identify and take ownership of risks related to electronic discovery, set policy relating to electronic discovery and achieve compliance with external and internal requirements relating to electronic discovery.

5.4 Overview of Part 3: Code of practice for electronic discovery

Part 3 identifies requirements and guidance associated with the electronic discovery processes and related measures.

5.5 Overview of Part 4: ICT readiness for electronic discovery

Part 4 provides guidance on the ways an organization can be better prepared to address electronic discovery from the perspective of both technology and processes.

6 Overview of electronic discovery

6.1 Background

Electronic discovery is increasingly important both within organizations and in some jurisdictions' legal systems, and this trend is expected to continue as more and more electronic records and information (or ESI) are created, modified, manipulated, used, and ultimately destroyed without ever taking on a physical form (e.g., a printed document). The emergence of ESI as the preferred representation of information is introducing new challenges for the legal community associated with locating the ESI, handling massive quantities of data, preservation and retention of ESI, authenticity, data integrity, data confidentiality, and data or media sanitization, etc. Failure to appropriately handle the electronic discovery processes as well as the ESI can result in costly rework, unnecessary costs, possible sanctions, and legal liabilities.

This International Standard addresses these challenges by:

- promoting a common approach, understanding, and language for electronic discovery
- encouraging practical and cost-effective discovery by those tasked with managing ESI through the process
- identifying competency areas for those involved in electronic discovery
- promoting the proactive use of technology, in reducing costs and risks, while increasing efficiencies throughout the discovery process
- suggesting ways of avoiding inadvertent disclosures of potentially privileged, confidential, or sensitive ESI

The overriding goal is to help organizations comply with electronic discovery objectives and obligation, if any.

6.2 Basic concepts

Parties need to consider in advance the following electronic discovery issues:

- Scope of electronic discovery;
- Establishing a team with responsibilities for each aspect of a typical electronic discovery project including identification of individuals responsible for ESI preservation;
- Identification of systems holding potentially relevant ESI;
- Identification of potentially relevant ESI;
- Anticipated costs and proposed allocation of same;
- Preservation of ESI, including the preservation notice process;
- Disclosure of the programs and manner in which ESI is stored;
- Collection/acquisition of ESI;
- Processing of ESI;

- 1 — Review and analysis of ESI;
- 2 — Production of ESI including the form of production.
- 3 Those engaging in electronic discovery will have many influencing factors specific to the context. Cost can be
- 4 significant among these. The primary cost drivers include:
- 5 — *Collection*: Finding and retrieving the potentially relevant ESI;
- 6 — *Volume*: The raw quantity of ESI to be processed or reviewed by human eyes;
- 7 — *Number of Custodians*: The number of sources involved in the collection of data can increase exponentially
- 8 the amount of time and effort involved;
- 9 — *Human Review*: The need for qualified people who can perform the review and recognize the specific ESI
- 10 meeting relevance, privilege, or other review criteria;
- 11 — *Case Complexity*: Simple cases may require a limited scope and review process, but more complex cases
- 12 can involve elaborate document review strategies and processes.

13 **6.3 Objectives of electronic discovery**

14 General electronic discovery objectives include:

- 15 — Address the differences in discovery between traditional forms of information and ESI;
- 16 — Identify potentially relevant sources of ESI;
- 17 — Properly preserve and retain potentially relevant ESI;
- 18 — Process that ESI into a format which will facilitate its efficient searching or review;
- 19 — Produce potentially relevant ESI in a form which is useable by the requesting party;
- 20 — Minimise the risk of missing potentially relevant ESI;
- 21 — Consider the proportionality of the response in the context of the matter and the costs;
- 22 — Utilise technology in order to reduce risks and costs throughout the project.

23 **6.4 Electronic discovery foundation for success**

24 **6.4.1 General**

25 Electronic discovery often involves parties with conflicting interests and in a worst case scenario they can be

26 adversarial parties. Electronic discovery can be key to resolving a conflict or matter, but only when it is conducted

27 on a foundation that facilitates a measure of trust.

28 For electronic discovery, this foundation includes adequately addressing competency, candour, cooperation,

29 completeness, and proportionality issues.

6.4.2 Competency

Given the complexities associated with electronic discovery, it is important that the individuals engaging in the electronic discovery process have the relevant technical and legal competencies. They may need to be able to demonstrate that they are properly trained and have sufficient technical and legal understanding to handle ESI appropriately and to execute the electronic discovery process on behalf of a party.

6.4.3 Candour

The parties conducting electronic discovery are expected to adhere to the highest standards of professionalism and ethical conduct. This means the parties have an obligation to correct and supplement the record (e.g., additional disclosures or to amend prior responses). In addition, purposeful sluggishness in executing the electronic discovery process needs to be avoided by all parties involved.

6.4.4 Cooperation

Cooperation on issues relating to the preservation, collection, search, review, and production of ESI is often expected in the courts, and further, such cooperation typically does not compromise representation of a client. Cooperation in reasonably limiting ESI discovery requests, on the one hand, and in reasonably responding to ESI discovery requests, on the other hand, tends to reduce costs and delay. Cooperative exchanges of information at the earliest possible stage of discovery are also particularly important.

6.4.5 Completeness

Within electronic discovery, tension exists in processes between completeness, on the one hand, and burden and cost, on the other. Responding parties seek to produce all responsive ESI, and at the same time they also seek to identify only the responsive ESI, in order to guard against overproduction or waiver of privilege.

6.4.6 Proportionality

With the explosive growth of ESI, there are increased concerns over how to best address the costs and burdens associated with the discovery process. One approach to address this problem is to take steps to help ensure that the benefits of discovery be commensurate with the corresponding burdens.

6.5 Planning and budgeting an electronic discovery project

The very drivers behind an electronic discovery project make it difficult to plan such a project many months in advance. As such, they are typically managed more like an incident response, which can increase costs significantly. Regardless of the urgency of the request, as with any project, time invested in planning at the outset typically saves significant time and costs later in the project. This is especially so, as many of the steps in a typical electronic discovery project are disproportionately expensive to repeat at a later stage. For example, if the production structure and format are not agreed upon in advance of the review, and families of documents are not marked consistently, then it can cause the review to have to be partially repeated.

An important early step is to establish an electronic discovery project team that, at a minimum, includes a project sponsor and manager from the business/organization, a project manager from the legal or investigative team, and a project manager from the IT perspective. This triangle of communication between the business/organization, the legal/investigators, and the IT team, is vital to a successful project.

Also, an early step is the establishment of an electronic discovery project plan, with as much detail as possible. As with any project plan, the electronic discovery plan needs to contain the project milestones (e.g. identification, preservation/collection, processing, review, analysis, production, and possibly presentation), the individual steps required at each stage, and the individual assignments at each step.

Given the costs involved in a typical electronic discovery project, preparing and monitoring a detailed budget from the outset is an important consideration. This budget needs to take into account the diverse disciplines on the electronic discovery team, and the fact that they can consist of internal and external counsel, as well as internal and external IT or electronic discovery consultants. It is important to include a budget for each phase of the plan, and in some cases, each step of each phase, so that it is possible to determine if the proposed approach is proportionate to the matter at hand.

7 Electronically Stored Information (ESI)

7.1 General

ESI is now an integral part of both business and individual environments. Consequently, it forms an increasingly important source of evidence in modern disputes or matters.

ESI needs to be considered at the earliest possible stage in a matter. It can be extremely fragile and is easily lost or modified, even though apparently inconsequential processes, such as opening a document. It is not necessary to undertake the full electronic discovery process from the outset; however completing the identification, preservation, and possibly the collection phases in the early days after becoming aware of a matter would be considered good practice and can lead to significant cost savings in the longer term.

Managing ESI increasingly impacts businesses and individuals; the volume, size, complexity, and range of ESI can often be overwhelming. ESI management is often not a priority until the true value and cost of locating ESI becomes apparent as part of a matter. Organizations and individuals frequently:

- Focus their ESI retention efforts on retention for purely business operational purposes rather than considering the wider context;
 - Have minimal consideration of their compliance obligations in respect of electronic records;
 - Have a limited understanding of the evidential value of good business records;
 - Do not have a good understanding of the costs and risks associated with poor information management practices.
- Poor ESI management can add challenges when it comes to identifying and retrieving ESI in response to a discovery or regulatory request because:
- ESI is often and unnecessarily stored beyond its required lifespan;
 - There is often little knowledge within the organization as to where potentially relevant ESI can be found;
 - The volume and complexity is overwhelming even for IT professionals;
 - Turnover of staff and organizational changes (e.g., mergers, acquisitions, and divestitures) which result in retention of ESI but the loss of organizational knowledge and context;
 - The IT environment and systems may be poorly documented.

These factors can lead to expensive and disruptive searches through vast quantities of ESI in order to locate that which may be relevant to the matter at hand. This can introduce delays and increase the cost of the discovery process, in addition to increasing the risk of potentially relevant ESI being overlooked.

The point of relevancy is important to note because it may be a determining factor in how ESI becomes potential digital evidence. In some jurisdictions, spoliation of potential digital evidence can have legal consequences, so additional care may be warranted when handling ESI in such jurisdictions.

7.2 Common types of ESI

7.2.1 General

Categorizing ESI sources as readily accessible (or 'active') or not-readily accessible (or, inactive, residual, or legacy), with a justification for each categorization is an important activity in the early stages of electronic discovery. In conjunction with budget preparation, this categorization will assist in determining the proportionality of preserving and collecting such sources.

7.2.2 Active data

This type of ESI is "actively" in use and resides on employees' computer hard drives or other storage devices and in the organization's servers, drives and databases. Active data generally can be accessed in a file manager or in the application in which it was created. Users can access it immediately without restoration or reconstruction. With the increasing popularity of cloud computing and Internet-based computing services, it may also reside on the storage devices of outside service providers. Most cases and investigations call primarily for the preservation and production of active files.

Active files may be relatively easy to access and collect, at least compared to other types of ESI. They can also be easily deleted or altered, thus preservation needs to be considered at the earliest possible time.

7.2.3 Inactive data

This type of ESI is related to closed, completed, or concluded activities, including ESI an organization maintains for long-term storage and record keeping purposes, but which is not immediately accessible to the user of a computer system. It may include many of the same sources of data described above in relation to active data.

Archived data is often stored in a compressed format and may be maintained on system drives or off-line devices, including backup tapes or disks and optical media. Some systems allow users to retrieve archival data directly, while other systems require the assistance of an IT professional. Challenges in preservation and collection include identifying relevant inactive and archived data, locating where and how it is stored, and restoring it from a compressed format.

Another form of inactive data is the ESI stored within data protection systems (e.g., backups). This inactive data can be a source of problems because it tends to be short-term (e.g., the backup media are rotated on a regular basis), there may not be any mechanism short of doing a full recovery to determine what is on the media, and stored data may only contain fragments (e.g., only the changes from the last backup). To complicate matters, IT personnel can make extra backups, which fall outside of normal operations (e.g., rotation, documentation, etc.) and it can be extremely difficult to identify these potential sources of ESI. This type of data has all the same challenges as archived data with the additional element of short retention periods, which requires quick action to suspend the automatic destruction of this ESI.

Clause 7.3.3.3 provides additional information on backups and archives.

7.2.4 Residual data

This type of ESI is hidden and cannot be viewed in applications (such as system files) or has been erased, fragmented, or damaged. Collecting this type of ESI usually requires a forensic copy—i.e., an exact, bit-by-bit copy of the entire physical storage media (e.g., hard drive, CD, DVD, tape), including all active and residual data and unallocated or slack space on the media.

Imaging and then extracting the residual data may require a digital evidence specialist (see ISO/IEC 27037) to operate special tools and can be time consuming and expensive. Making a bit-by-bit digital evidence copy may be unwarranted unless residual data is relevant and necessary in the matter. In some cases, however, companies may choose to image the hard drives of particularly important key custodians to ensure that all their data is preserved, including files that the custodian may have unintentionally, or intentionally, deleted or partially overwritten.

7.2.5 Legacy data

This type of ESI is created by software or hardware that is outmoded or has become obsolete (legacy systems). A legacy system may be one that the company still uses but that the hardware or software vendor no longer supports. Or, it may be a system that the company has decommissioned but retains in case its information is needed in the future.

The relevance of legacy data may be difficult to determine without restoration or reconstruction, and it may be costly to do so. In addition to preserving the legacy data itself, the company may need to retain the legacy hardware and software if there is no other way to view or use the data.

7.3 Common sources of ESI

7.3.1 General

Potentially relevant ESI in litigation and investigations can be found in a wide range of sources. To help identify these sources, it is important to consider systems and resources under the direct control and access of custodians as well as those that not under the control of custodians.

Custodian ESI sources are those sources of ESI which an individual custodian has direct custody over. These include sources such as the custodian's laptop or email mailbox. Non-custodian ESI sources are those which one or more custodians has access to, but it is likely that another custodian, such as an IT administrator has control over. These include centralised ESI sharing systems, such as databases, applications, and shared folders.

7.3.2 Custodian data sources

7.3.2.1 Communications

Email is often the central source of potentially relevant ESI in litigation and investigations. But other forms of communication are also prevalent and may also warrant consideration for preservation—instant messaging and chat, for example. It may be appropriate to consider whether ESI stored in an electronic fax system, copier, or in a videoconferencing system, is available and needs to be preserved. And, in some situations, audio recordings may exist that need to be considered for preservation.

7.3.2.2 Desktop, laptop or home computers

Potentially relevant ESI may be present on custodians' desktops, laptops or home computers. Even if an organization has network-based document management systems, employees may have also been given the ability to save documents on a local hard drive.

Although some organizations have policies prohibiting employees from using non-work issued computers, it may be advisable in certain situations to confirm whether, for example, key custodians in fact complied with the policy. Custodians may also have copied documents onto removable storage media, such as thumb drives, external hard drives, DVDs or CDs.

7.3.2.3 Mobile devices

Increasingly, mobile device such as mobile phones, smart phones, tablets, Global Positioning Systems (GPS), etc. generate and store ESI. These mobile devices can be important sources of potentially relevant ESI, especially when a custodian has been identified as potential party in a matter at hand.

7.3.3 Non-custodian data sources

7.3.3.1 Databases and applications

ESI related to dynamic databases may be relevant in some cases. For example, an employment matter may involve ESI from a company's human resources system, an antitrust matter may involve customer relationship management, sales or production systems, and a financial fraud case may involve accounting and finance systems. Depending on the issues, a matter may involve an organization's electronic document management systems (EDMS), electronic records management systems (ERMS), or collaborative tools.

Some database applications automatically purge data after a particular time period, so it can be advisable to identify and suspend such processes if necessary. Additionally, legacy systems may exist and their data considered for preservation.

7.3.3.2 Network storage

ISO/IEC 27040 describes Network Attached Storage (NAS) as a data storage technology that provides file-level access to heterogeneous clients over a network. NAS enables a file system physically residing on one server or device to be accessed by remote client computers, appearing to users as a local file system. NAS systems are typically designed and build specifically for NAS purposes, but general purpose server computers can also be used.

Documents may be stored in various places on an organization's internal network (e.g., shared drives, network disk drives, and servers). File servers and NAS are important sources of ESI because they are designed to provide centralized storage that can be easily shared and protected. Other forms of network storage include Web-based file services (e.g., WebDAV), cloud storage, and FTP servers.

A Storage Area Network (SAN) as described by ISO/IEC 27040 is a specialized, high-speed network that provides block-level network access to storage. SANs are typically composed of servers, switches, storage elements, and storage devices that are interconnected using a variety of technologies, topologies, and protocols. SANs can also span multiple sites. A SAN presents storage devices (such as disk arrays and tape libraries) to a server operating system such that to the server, the storage appears to be locally attached. This simplified presentation of storage to a server is accomplished through the use of different types of virtualization.

SANs can contain massive quantities of ESI, but this ESI can only be accessed by the systems and servers that are configured to connect/mount specific parts of the SAN. This means that each server that uses the SAN can be a source of ESI, but it is unlikely that the block-based storage can be a direct source of ESI.

7.3.3.3 Backups

It is common for organizations to back up the ESI on their information systems onto data protection systems such as tape or other media for disaster recovery and business continuity purposes. These backups typically have a relatively short shelf life, so the associated media is often recycled within 30-90 days (sometimes referred to as

rotating the backup media). In addition, the recovery process can be complex and cumbersome (e.g., appropriate recovery storage space has to be found, the backup may not be on a single piece of media¹⁾, etc.).

7.3.3.4 Electronic archives

An electronic or digital archive is a data repository that is typically part of a records management process that ensures protection, maintenance and accessibility of ESI, beginning from the moment of creating the ESI and ending with its disposition (i.e., destruction according to retention policies) or forever. ESI contained in archives are typically official business records, documents retained for compliance purposes, legacy documents (historical value), etc. The contents of such archives and their management are typically driven by organizational policies that are then implemented the records management system (e.g., destruction may take place automatically based on the expiration of a retention period).

7.3.3.5 Social media

Social media is ESI that is shared among groups of people mostly for social purposes, but increasingly it has been used for business purposes. It tends to reside outside of an organizations immediate control, resulting in some additional challenges in acquiring a copy of it. It is important that ESI is considered at the identification phase of the project and steps are taken early in the project to determine if it will be required, and if so, the steps required to preserve a copy of it. This determination needs to take into account the rapid pace by which it changes and can be made not-readily accessible.

7.3.4 Potentially excluded sources of ESI

Not all sources of ESI need to be preserved; the following sources of ESI generally are not discoverable in most matters:

- “deleted,” “slack,” or “unallocated” data on hard drives;
- random access memory (RAM) or other ephemeral data;
- on-line access data such as temporary internet files, history, cache, cookies, etc.;
- data in metadata fields that are frequently updated automatically, such as last-opened dates;

NOTE Careful consideration and consultation needs to be given to which metadata fields to be preserved, as it can be difficult, and frequently impossible to reverse once changed. For example, metadata information such as when a document was first created or last modified can be crucial in filtering ESI later in the process, so it needs to be maintained.

- backup data that is substantially duplicative of data that is more accessible elsewhere;
- other forms of ESI whose preservation requires extraordinary affirmative measures that are not utilized in the ordinary course of business.

It can be beneficial to attempt to reach an agreement with litigation opponents or investigators that such ESI does not need to be preserved.

¹⁾ Backups can take many forms, but the most common are full backups (all the data is captured on the media), differential backups (captures all changes made since the last full backup), and incremental (captures all changes made since the last backup). A backup solution may apply multiple of these approaches.

7.4 ESI representations

The ESI associated with a particular matter can include word processing files, spreadsheets, email, databases, drawings, photographs, data from proprietary applications, website data, voice mail, and much more. The collection and production formats for ESI files can be classified as native, near-native, image (near-paper) and paper.

7.4.1 Native formats

Files in the format they were created and maintained are known as a native files. Native format is often recommended for files that were not created for printing, such as spreadsheets and small databases. For some file types the native format may be the only way to adequately produce the documents.

In most cases, this is the most efficient way to produce ESI, as it does not require the producing party to incur the cost of converting it to a different format; for the receiving party, it can require native application or provision of client's proprietary software to open files.

In the event that a party chooses to convert ESI into a different format, steps to ensure that elements of that ESI, such as metadata, are not unintentionally lost or obscured in the process may be necessary.

7.4.2 Near native formats

Some files (e.g., email and databases) cannot be reviewed without some form of conversion. For example, most email files have to be extracted and converted into individual files, and as a result, the original format is altered and they are no longer in native format. There is no standard format for near-native files, but email files are typically converted to a structured text format such as HyperText Markup Language (HTML or files with a .htm or .html extension) or Extensible Markup Language (XML or files with a .xml extension); these formats do not require special software for viewing.

Large databases and data compilations are commonly produced in near-native format. Databases can comprise massive amounts of completely undifferentiated tables of data. Enterprise business systems may contain hundreds of tables and thousands of fields of data. The systems may require various database platforms and proprietary software. For these reasons, large databases and data compilations are generally not produced in native format. These databases often need to be analysed by appropriate personnel to identify the responsive data and determine the appropriate near native format.

Exports from these databases are often produced as text delimited files. In some cases text files are produced with a database diagram, data dictionary, metadata or software. Data may also be exported to common spreadsheet formats.

7.4.3 Image (near paper) formats

ESI can also be produced in an image, or near paper, format. Rendering an image is the process of converting ESI or scanning paper into a non-editable digital file. During this process a "picture" is taken of the file as it exists or would exist in paper format. Based on the print settings in the document, the printer or the computer, data can be altered or missing from the image. Expertise in the field of electronic discovery and image rendering tools are necessary to minimize this risk.

7.4.4 Paper

Paper is used as paper or ESI is printed to paper and the paper is what is used. As with converting to image (see clause 7.4.3), printing documents to paper can result in missed or altered data. When ESI is converted to paper, it is recommended to use someone with expertise in the field of electronic discovery and image rendering tools to minimize this risk during the printing or image rendering process.

7.5 Non-ESI as part of discovery

While most business information is stored in electronic format, most discovery projects will involve at least small element of traditional hardcopy or paper documents. These will typically be lower in volume than their electronic equivalents, and most importantly, the process of identifying, preserving, and collecting them will in itself result in a large part of the processing phase being completed up front. This is due to the fact that it's more likely that the hardcopy documents which are collected will have been collected with the matter in mind, and therefore will likely be relevant to the matter.

With a focused set of hardcopy documents, it is typically only necessary to have them scanned into electronic format and then included in the review process alongside the electronic documents. Quite often, it will not be necessary to have a first pass (or relevancy) review completed on such documents, and they can typically be moved directly to the detailed second pass review. They are also an example of documents which are unlikely to have associated family members, and therefore consideration of families of hardcopy documents is often not required from an electronic discovery perspective. Such documents will however have the requirement to be reviewed, marked, and maybe redacted, before being included in a schedule alongside the electronic documents at the production phase. This makes it most efficient to use the same technologies and processes to manage hardcopy documents as is used for electronic documents.

There are two technical elements of the scanning process which are useful to consider. The first is that of scanning the hardcopy document to a searchable electronic format using a process known as Optical Character Recognition or OCR. This is not an exact science, as it is essentially reliant on the computer recognising text. Therefore, it may not result in every piece of text in the hardcopy document being recognised and made searchable in the electronic copy (this is especially so in the case of handwritten text). However, once the accuracy of the process is understood, and adequate QA procedures are put in place, then it may be possible to rely on searching such documents using electronic tools.

The second technical concept is that which is often referred to as coding. Electronic documents have metadata built into them. For example, an email will have the author, the recipient, and the date, along with the fact that it is an email. The technology tools used in the electronic discovery process automatically recognise these metadata fields and present them to the reviewer as such (i.e. the computer recognises an email). A scanned copy of a hardcopy document is however akin to a photocopy of the document, and as such does not contain metadata as to who wrote the document, when it was written, or who it was posted to (although this information may be in the content of the document itself). While there are some technologies available which can identify such metadata from scanned hardcopy documents, it is a difficult task, as the information required is not often in the same location (as it would be in an email, thus allowing the computer to identify and present it automatically). The solution to this is to have a human 'code' the documents. This essentially involves the document being scanned and OCR'd, and it is then passed to a 'coder', who will review the document (to the extent necessary) and take note of who sent it, who it was addressed to, and what date it was posted. This requires a great deal of human interpretation, and as such is quite a manual exercise. The output of this process is that the 'coding' information can be used to facilitate the processing and review phases. For example, it may be necessary to have one reviewer review all letters between two parties in chronological order. This 'coding' information can be used to identify all letters and then sort them by date; a task would not have been possible with just the scanned copies of the hardcopy documents alone.

8 Electronic discovery process

8.1 Overview

A typical electronic discovery project is conducted in a series of phases. It is however, usual that an overall electronic discovery project will be iterative in nature, with some phases being repeated as the matter evolves. In addition, not all projects will use all phases and all phases may not be carried out by the same team. Typical phases include:

- 1 — Identification - To identify the possible sources of ESI which may, by its contents or as a material object, be
2 tendered as evidence relevant to the facts in issue, and if agreed, the categories of discovery;
- 3 — Preservation - To preserve the ESI sources identified;
- 4 — Collection - To make a copy of the ESI sources, while preserving the integrity of the ESI;
- 5 — Processing - In the processing phase, ESI which has previously been collected is prepared for searching,
6 review, or other analysis, while preserving the integrity of the ESI;
- 7 — Review - The review phase is primarily concerned with performing a review of the documents resulting from
8 the filtering process undertaken in the processing phase in order to determine if they are responsive;
- 9 — Analysis - To take a deeper look at a document, for example to determine its provenance. This phase is not
10 always necessary;
- 11 — Production – To produce the ESI which have been determined as responsive from the review phase. These
12 are produced to the requesting party.

13 A detailed overview of each phase can be seen within this clause 8.

14 **8.2 Identification phase**

15 Typically, identification will involve, firstly determining what ESI exists, and secondly, identifying its location or the
16 means of accessing it.

17 Gathering information on the existence and location of potentially discoverable ESI is therefore a prerequisite to
18 analysing whether ESI needs to be preserved, collected, and reviewed, since a party clearly cannot produce ESI if
19 it does not know where its own documents are stored. Identification typically takes into consideration the facts of
20 the matter, preservation demands, disclosure requirements, and discovery demands, including categories of ESI
21 requested.

22 At a bare minimum, identification for electronic discovery involves identifying 1) key players, custodians, locations
23 of data and traceability of data to individuals and departments, and 2) key people involved in the discovery project
24 management and enforcement of the litigation hold, such as corporate legal counsel, personnel from IT and
25 records management departments, outside counsel and discovery consultants and service providers. The
26 identification strategy is paramount.

27 In some jurisdictions, courts, legislatures, and government regulators have developed rules concerning how
28 organizations identify ESI, particularly for purposes of civil and criminal proceedings, investigations, and audits. In
29 such jurisdictions, organizations may have a duty to take reasonable steps to identify and preserve potentially
30 relevant ESI when a litigation or investigation is reasonably anticipated or pending against them. Underlying this
31 duty to identify and preserve ESI, an organization has to be able to locate and preserve potentially relevant ESI in
32 a timely manner. Further, organizations may be expected to develop appropriate ESI management protocols and
33 be compliant with those rules. The Offices of the General Counsel and Information Technology (IT) are typically
34 held accountable for the enterprise-wide ESI management practices and generally depend upon the support of
35 leadership in other areas of the organization to effectively execute these practices. These stakeholders are
36 expected to take reasonable steps to ensure that all data potentially relevant to a matter or inquiry are identifiable.
37 As such, developing a strategy and executing upon a defined, defensible process for identification is critical.

38 Organizations need to have the ability to identify the particular systems that are likely to be subject to preservation
39 and disclosure requirements in investigations and legal proceedings. Being proactive and gathering timely
40 information about high risk systems will enable them to meet the expectations of the courts and regulators.
41 Organizations need to take steps to expand their current understanding of their information systems to include

what systems are in operation, how those systems or applications are related, where those systems are housed, which organizations those systems support, who key contact personnel are for those systems, and other information which could potentially be necessary for timely and adequate preservation and disclosure in litigations and investigations.

The typical corporate infrastructure includes a broad population of users representing the organization's everyday routine. Not everyone in a company may be relevant to a particular matter, but the IT organization, which is designed to support an entire infrastructure, will be. In many cases, key players need to be identified by opposing parties. Of course, the requesting parties don't want to limit a search that may result in responsive data being missed. It has been traditional to have the producing party decide who the key players are and target them for specific collection. The meet-and-confer conference held early on in the discovery process can be used to identify who the key players are and what kind of data can be expected from each.

Review the relevant pleadings and discovery requests to determine the relevant time period to the matter. Use these dates to assist in locating and culling relevant data.

Documentation of the identification process is essential throughout the case when questions arise regarding additional sources of information. More importantly, it can be used to demonstrate the identification process was defensible if it come under question.

8.3 Preservation phase

Preservation is necessary where potentially discoverable ESI may be lost or altered in the normal course of business before a party has the opportunity to collect it. It is prudent to preserve sources of potentially discoverable ESI as early as possible.

NOTE Failure to take adequate steps to preserve ESI can result in spoliation, which can negatively impact the organization in the matter.

Once a decision has been made that a duty to preserve has been triggered, the scope of that duty needs to be evaluated; decisions as to scope may address time frames, custodians, subject matter, and responsive information by source or system, category, or type. Considerations can include: the facts upon which the triggering event is based and the subject matter of the triggering event; whether the ESI is relevant to that event; the expense and burden incurred in preserving the ESI; and whether the loss of the ESI would be relevant to the matter.

As part of preservation, issue a legal hold or preservation notice to appropriate parts of the organization, including data custodians (if known) and IT administrators. If the ESI is held by third party, it may be appropriate to notify them as well.

8.4 Collection phase

The objective of collection is to take a copy of the agreed ESI sources so that their content can be processed and made available for review. One of the additional objectives of collecting ESI, in many cases, is to secure a forensically sound copy of certain ESI as it was stored on a particular date and time. This may be necessary if the admissibility or validity of the ESI is later questioned.

It is important that potentially relevant ESI and its associated metadata be collected in a manner that is defensible, targeted, proportionate to the matter, auditable and efficient. Collection methodologies that organizations may consider include employee self-collection, IT-assisted collection and collection by an outside service provider (e.g., an electronic discovery vendor). In addition, a variety of tools and approaches can be used to actually collect the target ESI (e.g., forensic imaging, making digital copies, data exports, etc.).

The method of collection may need to take into account the skills and tools available to those completing the collection, whether metadata needs to be preserved, and whether a chain of custody needs to be maintained. For

example, an untrained custodian completing a self-collection may be appropriate where the ESI is only emails, and thus less likely to have metadata accidentally altered. Whereas, it would unlikely be appropriate for a custodian to complete a forensic collection where they are the subject of allegations of fraud.

NOTE ISO/IEC 27037 differentiates between collection and acquisition. Collection within this International Standard is more closely aligned with acquisition in ISO/IEC 27037.

8.5 Processing phase

In order to efficiently search and review ESI, it is necessary to prepare, or process, the ESI. The extent and nature of processing required in any given project will depend on the nature of the ESI collected, the technology being used, and the expected review process.

Some primary goals of processing are to discern at an item-level exactly what data is contained in the universe submitted; to record all item-level metadata as it existed prior to processing; and to enable defensible reduction of data by “selecting” only appropriate items to move forward to review. All of this needs to happen with strict adherence to process auditing; quality control; analysis and validation, and chain of custody considerations.

Data may arrive at the processing stage in various formats which then need to be restored before subsequent work can be done (tapes, backups, etc.); individual files and email may need to be extracted from container files (PST, NSF, zip, rar, etc.); and certain types of data may need to be converted to facilitate further processing (legacy mail formats; legacy file formats). During these processing stages individual items are catalogued and their associated metadata is captured.

Rarely is it necessary to review all items that are submitted for processing. A number of data reduction opportunities are usually available. Processing is further broken down to five sub-processes, namely: Assess Data / Plan, Prepare Data, Select and Normalise Data, Validate Output / Exception Handling, and Prepare Output and Export. Assessment may allow for a determination that certain data need not move forward; Preparation involves performing activities against the data which will later allow for specific item-level selection to occur (extraction, indexing, hashing, etc.); Selection involves de-duplication; searching; and analytical methods for choosing specific items which will be moved forward; Output allows for transport of reviewable items to the next phases of electronic discovery.

8.6 Review phase

As with the traditional discovery process, it is necessary to review a document in order to determine whether any form of privilege applies, to assess relevance, and often to categorise it according to issues in a matter or based on a specific schedule of requests. Electronic documents or ESI, are no different in their requirement for review. The main difference is that there are usually more electronic documents than hardcopy (paper) documents. However, the technology used in the processing phase is used to reduce the number of documents for review, and to focus on those most likely to be relevant.

While much of the work undertaken in order to complete the first five phases will be carried out by IT or electronic discovery specialists, in conjunction with legal advisors, review work is typically carried out by subject matter experts, such as trained legal professionals, accountants, or investigators.

It is rare that any two reviews will be identical from a process perspective (they should probably never be identical from a document content perspective). The approach taken to a review will vary, depending on the factors of the case, the type of data being reviewed, and the objectives of the review. Most reviews will however have a number of common factors, therefore it is prudent to have a suggested starting point for a generic review, and then amend as is required for each individual set of circumstances.

8.7 Analysis phase

The objective of analysis is to take a deeper look at a document, for example, to determine its provenance. Structured data, such as accounting systems, can also be analysed to generate insights into specific transactions, or patterns of transactions. The key difference between the review phase and the analysis phase is that the review phase is typically focused on determining if a document is relevant to a matter, while the analysis phase is not so much focused on the content of a document, but its provenance, or its lifecycle.

NOTE While the analysis phase follows the review phase in this International Standard, analysis can be deployed in many phases of discovery as well as pre-discovery.

The analysis phase can be further sub-divided to content analysis, which concerned with understanding the circumstances, facts and potential digital evidence in a litigation or investigation, and process analysis, which is concerned with understanding the efficacy of the methods employed during discovery and the decisions reached based on analysis.

8.8 Production phase

Quite often the final step in the discovery process is providing a copy of the ESI (and hardcopy documents) which have been found to be relevant to the requesting party. It is vital that parties engage early in the process, so that what is produced at the end of the process is not a surprise.

Requests for ESI from litigants and third-parties for ESI are frequently met with objections that the requests are burdensome and overly broad. In addition, in electronic discovery, technical, highly complex issues may render requests inherently ambiguous and render compliance very difficult. To avoid, or contain, potential problems arising as a result of these issues, document requests and subpoenas for the production of ESI, and objections to those requests and subpoenas for ESI, need to be written in plain, clear language with as much specificity as possible under the circumstances. See also clause 6.4.

The output of the review phase would be a number of documents which have been marked as relevant at the second pass review, some of which will also be marked as privileged, or partly-privileged, and may have redactions applied.

In regards to the production format, there are a vast number of technical options available, making it vital to engage with the receiving party early in the process in order to gain agreement. Each jurisdiction will also have its own requirements. One common and cost effective option is that documents which have been marked as relevant are produced in their native format, along with a schedule listing their metadata details, categories, etc. (in a 'load file'). Documents which have been redacted are produced alongside their non-redacted counterparts, however would be in a redacted (e.g. PDF/TIFF) format. Documents which have been marked as privileged would not be produced, however a schedule of such documents might be produced.

9 Additional considerations

9.1 Information management

There are several major reasons why organizations face significant challenges when it comes to identifying and retrieving ESI in response to a discovery or regulatory request:

- It is often and unnecessarily stored beyond its required lifespan;
- There is often little knowledge within the organization as to where potentially relevant ESI can be found;
- The volume and complexity is overwhelming even for Information Technology (IT) professionals;

— Turnover of staff and organizational changes (e.g., mergers, acquisitions, and divestitures) which result in retention of ESI but the loss of organizational knowledge and context;

— The IT environment and systems may be poorly documented.

These factors lead to unavoidable, expensive, and disruptive trawls through vast quantities of ESI in order to locate that which may be relevant to the matter in hand. This can introduce delays and increase the cost of the discovery process, in addition to increasing the risk of potentially relevant ESI being overlooked.

ISO/IEC 27050-4 addresses many of these issues within the context of electronic discovery.

9.2 ESI presentation

While not considered a phase of the electronic discovery process, it is important to understand how ESI will ultimately be used in a matter (e.g., presented in court).

The presentation of ESI can be a challenge for attorneys and paralegals. In the past, exhibits were presented in paper form and still are in many cases today. Technology has developed over the last decade making it easier to present exhibits in near-paper or “image” format. Due to the nature of electronically stored information and the advent of native and near native document productions, some cases now require the legal team to present exhibits in native format.

9.3 Chain of custody and provenance

Depending on the matter at hand, it may be important to track or determine information regarding the creation, modification history, influences, ownership, or other provenance or lineage information associated with ESI. Some of this information may be contained in metadata or it may be generated as part of the electronic discovery process. This provenance information can be essential for making informed judgements about ESI quality, integrity, and authenticity.

In some cases, provenance information is not sufficient to demonstrate quality, integrity, and authenticity. In such cases (e.g., criminal investigations or prosecutions), formal chronological documentation that shows the custody, control, transfer, and disposition of ESI (more likely, potential digital evidence) is necessary. It is important to recognize when proof of chain of custody is needed and to ensure that the requirements are met.

9.4 Protection of ESI

9.4.1 General

Throughout the electronic discovery process, the parties involved in a matter are gathering, handling, and manipulating ESI. Often this ESI has been extracted from a computing or storage environment that is specifically designed to protect it. Similar protections may be needed for the ESI that has been removed or copied from these environments.

9.4.2 Long-term retention of ESI

Electronic discovery is commonly employed early in litigation, audit, government investigation or other such matters. While the matter proceeds, the parties need to retain the associated ESI in such a way that it continues to be available and its integrity is maintained.

It is important to consider the timeframes involved when making decisions about long-term retention of ESI. There are significant differences between the approaches for retaining ESI for a few weeks or months versus retaining ESI for decades (e.g., complex civil litigation that goes through multiple appeals) in electronic archives.

1 An additional consideration is whether data protection and privacy requirements affect how long personal data
2 may be retained and if the matter requires normal data retention periods to be suspended. This will vary
3 significantly between jurisdictions.

4 **9.4.3 Maintaining ESI confidentiality**

5 ESI often contains proprietary, privileged, and sensitive information that needs to be handled and stored in a way
6 that protects the confidentiality of the information. Failure to adequately control sensitive ESI can result in serious
7 repercussions if there is a data breach.

8 ISO/IEC 27001, ISO/IEC 27002, and ISO/IEC 27040 provide relevant guidance on security controls that can help
9 an organization devise an adequate approach to protecting sensitive ESI. These controls typically include access
10 control, encryption, and key management.

11 **9.4.4 Disposition of ESI**

12 When ESI is no longer needed it is important to dispose of it in a way that avoids data breaches. This typically
13 means that the logical storage or the storage media used to retain the ESI has to be properly sanitized (e.g.,
14 cleared using overwrite techniques or cryptographic erase).

15 ISO/IEC 27001, ISO/IEC 27002, and ISO/IEC 27040 provide relevant guidance on sanitization controls that can
16 help an organization devise an adequate approach to guard against media-based data breaches.

Annex A (informative)

Potentially relevant standards

A.1 General

This International Standard describes part of a comprehensive investigative process which includes, but is not limited to, the following topic areas:

- Incident management, including preparation and planning for investigations;
- Handling of digital evidence;
- Use of, and issues caused by, redaction;
- Intrusion prevention and detection systems, including information which can be obtained from them;
- Security of storage, including sanitization of storage;
- Ensuring that investigative methods are fit for purpose;
- Carrying out analysis and interpretation of digital evidence;
- Understanding principles of digital evidence investigations;
- Security incident event management, including derivation of evidence from systems involved in this;
- The relationship between electronic discovery and other investigative methods, as well as the use of electronic discovery techniques in other investigations;
- Governance of investigations, including forensic investigations.

These topic areas are addressed, in part, by the following ISO/IEC standards:

- ISO/IEC 27037: Guidelines for the identification, collection, acquisition and preservation of digital evidence.

This International Standard describes the means by which those involved in the early stages of an investigation, including initial response, can assure that sufficient potential digital evidence is captured to allow the investigation to proceed appropriately.

- ISO/IEC 27038: Specification for digital redaction.

Some documents can contain information that must not be disclosed to some communities. Modified documents can be released to these communities after an appropriate processing of the original document. The process of removing information that is not to be disclosed is called “redaction”.

The digital redaction of documents is a relatively new area of document management practice, raising unique issues and potential risks. Where digital documents are redacted, removed information must not be

recoverable. Hence, care needs to be taken so that redacted information is permanently removed from the digital document (e.g. it must not be simply hidden within non-displayable portions of the document).

ISO/IEC 27038 specifies methods for digital redaction of digital documents. It also specifies requirements for software that can be used for redaction.

— ISO/IEC 27040:—: Storage security.

This International Standard provides detailed technical guidance on how organizations may define an appropriate level of risk mitigation by employing a well-proven and consistent approach to the planning, design, documentation and implementation of data storage security. Storage security applies to the protection (security) of information where it is stored and to the security of the information being transferred across the communication links associated with storage. Storage security includes the security of devices and media, the security of management activities related to the devices and media, the security of applications and services, and security relevant to end-users during the lifetime of devices and media and after end of use.

Security mechanisms like encryption and sanitization can affect one's ability to investigate by introducing obfuscation mechanisms. They need to be considered prior to and during the conduct of an investigation. They can also be important in ensuring that storage of evidential material during and after an investigation is adequately prepared and secured.

— ISO/IEC 27041:—: Guidance on assuring the suitability and adequacy of incident investigation methods.

It is important that methods and processes deployed during an investigation can be shown to be appropriate. This document provides guidance on how to provide assurance that methods and processes meet the requirements of the investigation and have been appropriately tested.

— ISO/IEC 27042:—: Guidelines for the analysis and interpretation of digital evidence.

This International Standard describes how methods and processes to be used during an investigation can be designed and implemented in order to allow correct evaluation of potential digital evidence, interpretation of digital evidence and effective reporting of findings.

The following ISO/IEC projects also address, in part, the topic areas identified above, and may lead to the publication of relevant standards at some time after the publications of this standard:

— ISO/IEC 27035: Information security incident management
[Status: at the time of writing, all parts of this International Standard were in development]

This is a three part standard that provides organizations with a structured and planned approach to the management of security incident management. It is composed of

— ISO/IEC 27035 Part 1: Principles of incident management.

This part presents basic concepts and phases of information security incident management. It combines these concepts with principles in a structured approach to detecting, reporting, assessing, responding and applying lessons learned.

— ISO/IEC 27035 Part 2: Guidelines to plan and prepare for incident response.

This part presents the concepts to plan and prepare for incident response. The concepts, including incident management policy and plan, incident response team establishment and awareness briefing and training, are based on the plan and prepare phase of the model presented in ISO/IEC 27035-1. This Part also covers the "Lessons Learned" phase of the model.

— ISO/IEC 27035 Part 3: Guidelines for incident response operations.

This part includes staff responsibilities and practical incident response activities across the organization. Particular focus is given to the incident response team activities such including monitoring, detection, analysis, and response activities for the collected data or security events.

- ISO/IEC 27044: Guidelines for Security Information and Event Management (SIEM).
[Status: at the time of writing, this International Standard was in development]

This provides guidelines to organizations in preparing to deploy Security Information & Event Management Processes/Systems. In particular, it addresses the selection, deployment and operations of SIEM. It intends specifically to offer assistance in satisfying requirements of ISO/IEC 27001:2005: regarding the implementation of procedures and other controls capable of enabling prompt detection and response to security incidents, to execute monitoring and review procedures to properly identify attempted and successful security breaches and incidents.

- ISO/IEC 30121: Governance of digital forensic risk framework.
[Status: at the time of writing, this International Standard was in development]

This International standard provides a framework for Governing bodies of organizations (including owners, board members, directors, partners, senior executives, or similar) on the best way to prepare an organization for digital investigations before they occur. This standard applies to the development of strategic processes (and decisions) relating to the retention, availability, access and cost effectiveness of digital evidence disclosure. This International Standard is applicable to all types and sizes of organizations. The International Standard is about the prudent strategic preparation for digital investigation of an organization. Forensic readiness assures that an organization has made the appropriate and relevant strategic preparation for accepting potential events of an evidential nature. Actions may occur as the result of inevitable security breaches, fraud, and reputation assertion. In every situation Information Technology (IT) needs to be strategically deployed to maximize the effectiveness of evidential availability, accessibility and cost efficiency

A.2 Issues with investigative concepts and terminology

Within the investigative-oriented International Standards identified in A.1, defined terminology and identified processes are either in conflict with electronic discovery concepts and terminology or can further complicate an already challenging activity. This clause A.2 identifies some of the issues that are likely to cause confusion and attempts to offer some clarity.

From a terminology perspective, the following terms have been defined:

- **Digital evidence (ISO/IEC 27037):** information or data, stored or transmitted in binary form, that may be relied on as evidence; data or information of any kind and from any source, whose temporal existence is evidenced by being stored in or on any electronic medium;
- **Potential digital evidence (ISO/IEC 27042):** digital evidence that has not yet been admitted as such in a court of law or other legal proceeding;
- **Legal digital evidence (ISO/IEC 27042):** digital evidence which has been accepted into a judicial process;
- **Electronically Stored Information (ISO/IEC 27040):** data or information of any kind and from any source, whose temporal existence is evidenced by being stored in or on any electronic medium.

ISO/IEC 27050 focusses on ESI, which is much broader than either digital data or digital evidence. That said, it is possible that a successful execution of electronic discovery can result in distilling ESI into various forms of digital evidence, depending on the nature of the matter.

Another potential problem area includes the following process elements:

- **Analysis (ISO/IEC 27042):** evaluation of potential digital evidence in order to assess its relevance to the investigation;
- **Analysis (ISO/IEC 27050-1):** element of an electronic discovery process focused on evaluating Electronically Stored Information for content and context to identify facts, relationships, key patterns, and other features that can lead to improved understanding of an ESI corpus;
- **Acquisition (ISO/IEC 27037):** process of creating a copy of data within a defined set;
- **Collection (ISO/IEC 27037):** process of gathering the physical items that contain potential digital evidence;
- **Collection (ISO/IEC 27050-1):** element of an electronic discovery process focused on gathering Electronically Stored Information and other related material;
- **Identification (ISO/IEC 27037):** process involving the search for, recognition and documentation of potential digital evidence;
- **Identification (ISO/IEC 27050-1):** element of an electronic discovery process focused on locating potential sources and the criteria for selecting potentially relevant Electronically Stored Information;
- **Preservation (ISO/IEC 27037):** process to maintain and safeguard the integrity and/or original condition of the potential digital evidence;
- **Preservation (ISO/IEC 27050-1):** element of an electronic discovery process focused on ensuring that Electronically Stored Information is protected against inappropriate alteration or destruction.

The ISO/IEC 27037 and ISO/IEC 27042 terminology tends to be aligned with forensics. Unfortunately, forensics and electronic discovery use the same terminology, but for different things. As a case in point, collection within the context of forensics means grabbing the physical systems and acquisition is making copies of the data; however, collection from the an electronic discovery perspective is actually closer to forensic-oriented collection.

Bibliography

- [01] ISO Guide 73:2009, Risk management – Vocabulary
- [02] ISO/IEC FDIS 27040, *Information technology — Security techniques — Storage security*
- [03] ISO/IEC 27037:2012, *Information technology — Security techniques — Guidelines for identification, collection, acquisition, and preservation of digital evidence*
- [04] ISO/IEC DIS 27041, *Information technology — Security techniques — Guidance on assuring suitability and adequacy of investigation methods*
- [05] ISO/IEC DIS 27042, *Information technology — Security techniques — Guidelines for the analysis & interpretation of digital evidence*
- [06] ISO/IEC DIS 27043, *Information technology — Security techniques — Investigation principles and processes*
- [07] ISO 15489:2001, *Information and documentation — Records management*
- [08] Electronic Discovery Reference Model (EDRM), <http://www.edrm.net>
- [09] Good practice guide to Electronic Discovery in Ireland, Version 1.0, 16 April 2013
- [10] New York Bar Association, *Best Practices in E-Discovery in New York State and Federal Courts*, Version 2.0, December 2012
- [11] *Seventh Circuit Electronic Discovery Pilot Program – Final Report on Phase Two*, May 2012, <http://www.discoverypilot.com/sites/default/files/Phase-Two-Final-Report-Appendix.pdf>