DRAFT INTERNATIONAL STANDARD ISO/IEC DIS 27038

ISO/IEC JTC **1** Secretariat: **ANSI**

Voting begins on
2013-01-14

Voting terminates on
2013-04-14

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION • МЕЖДУНАРОДНАЯ ОРГАНИЗАЦИЯ ПО СТАНДАРТИЗАЦИИ • ORGANISATION INTERNATIONALE DE NORMALISATION
INTERNATIONAL ELECTROTECHNICAL COMMISSION • МЕЖДУНАРОДНАЯ ЭЛЕКТРОТЕХНИЧЕСКАЯ КОММИСИЯ • COMMISSION ÉLECTROTECHNIQUE INTERNATIONALE

Information technology — Security techniques — Specification for Digital Redaction

ICS 35.040

THIS DOCUMENT IS A DRAFT CIRCULATED FOR COMMENT AND APPROVAL. IT IS THEREFORE SUBJECT TO CHANGE AND MAY NOT BE REFERRED TO AS AN INTERNATIONAL STANDARD UNTIL PUBLISHED AS SUCH.

IN ADDITION TO THEIR EVALUATION AS BEING ACCEPTABLE FOR INDUSTRIAL, TECHNOLOGICAL, COMMERCIAL AND USER PURPOSES, DRAFT INTERNATIONAL STANDARDS MAY ON OCCASION HAVE TO BE CONSIDERED IN THE LIGHT OF THEIR POTENTIAL TO BECOME STANDARDS TO WHICH REFERENCE MAY BE MADE IN NATIONAL REGULATIONS.

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.

Contents

Foreword

ISO (the International Organization for Standardization) and IEC (the International Electrotechnical Commission) form the specialized system for worldwide standardization. National bodies that are members of ISO or IEC participate in the development of International Standards through technical committees established by the respective organization to deal with particular fields of technical activity. ISO and IEC technical committees collaborate in fields of mutual interest. Other international organizations, governmental and non-governmental, in liaison with ISO and IEC, also take part in the work. In the field of information technology, ISO and IEC have established a joint technical committee, ISO/IEC JTC 1.

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

The main task of the joint technical committee is to prepare International Standards. Draft International Standards adopted by the joint technical committee are circulated to national bodies for voting. Publication as an International Standard requires approval by at least 75 % of the national 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 and IEC shall not be held responsible for identifying any or all such patent rights.

ISO/IEC 27038 was prepared by Joint Technical Committee ISO/IEC JTC 1, *Information technology*, Subcommittee SC 27, *Security techniques*.

Introduction

Some documents may contain information that must not be disclosed to some communities. Modified documents may be released to these communities after an appropriate processing of the original document. This processing may include the removal of sections, paragraphs or sentences with, where appropriate, the mention that they have been removed. This process is called the "redaction" of the document.

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).

This international standard specifies methods for digital redaction of digital documents.

Redaction may also involve the removal of document metadata or the removal of some information (e.g. an image) imported into the document.

It may be possible to identify redacted information in a redacted digital document by context. For example, the length of the redaction replacement text may indicate the length of the redacted information, and thus the information itself. This international standard introduces two levels of redaction:

— BASIC redaction where context is not taken into consideration;

— ENHANCED redaction where context is taken into consideration.

Redaction techniques can be used for the anonymization of the information in a document, for example by the removal of some names within sentences. It can also involve the removal of numbers within sentences and their replacement by "XXX"

Information technology — Security techniques — Specification for digital redaction

1 Scope

This international standard specifies characteristics of techniques for performing digital redaction on digital documents.   This international standard also specifies requirements for software redaction tools and methods of testing that digital redaction has been securely completed.

This international standard does not include the redaction of information from databases.

2 Terms and definitions

For the purposes of this document, the following terms and definitions apply.

2.1
anonymization
process by which personally identifiable information (PII) is irreversibly altered in such a way that a PII principal can no longer be identified directly or indirectly, either by the PII controller alone or in collaboration with any other party
[ISO/IEC 29100]

2.2
document
recorded information which can be treated as a unit
NOTE: Documents may contain text, pictures, video and audio content, metadata and other associated content

2.3
personally identifiable information
PII
any information that (a) can be used to identify the PII principal to whom such information relates, or (b) is or might be directly or indirectly linked to a PII principal
[ISO/IEC 29100]

2.4
redaction
permanent removal of information within a document

2.5
reviewer
individual(s) who assesses a document for redaction requirements
NOTE: There could be a series of individuals who assess a particular document

3 Symbols and abbreviated terms

For the purposes of this document, the following abbreviations apply.

PII   Personally Identifiable Information

PDF   Portable Document Format

OCR   Optical Character Recognition

XML   Extensible Markup Language

4   General principles of digital redaction

4.1   Introduction

Redaction is carried out in order to permanently remove particular information from a copy of a document. It should be used when, for example, one or two individual words, a sentence or paragraph, an image, a name, address and/or signature needs to be removed from a document prior to it being disclosed to individuals who are not authorized to view the removed information.

The process of digital redaction is not simply to remove information but also to indicate where necessary that some information has been removed, so that the reader knows that the document has been redacted. For example, there may be a need to know that some words or some paragraphs have been deleted in order to maintain the semantics of the non-redacted information.

4.2   Anonymization

One of the purposes of redaction is to remove personally identifiable information (PII) from a document (anonymization). Where such a purpose is applicable, then redaction processes shall be so designed such that the identity of the individual about whose information is being redacted is protected.

It may be, for example, that even though a name has been redacted from a document, the identity of the individual is evident from the remaining information. Where anonymization is required, all information that could be used to identify the individual shall be redacted. This shall include all information that could be used in conjunction with other information (which may be obtained from other sources) to identify the individual.
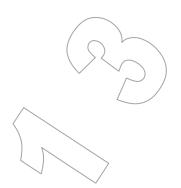
5   Requirements

5.1   Overview

Organizations should have the capability to identify documents that need to be redacted prior to their release to other parts of the organization or to others (such as the public).

Redaction should be performed or overseen by reviewers that are knowledgeable about the documents and can determine what information is to be redacted. If reviewers identifying such information do not carry out redaction themselves, their instructions shall be specific e.g. 'Memo dated …, paragraph no…, line starting… and ending…' etc.

Redaction shall be carried out on copies of the digital document. The redaction process shall result in the creation of a new digital document where complete and irreversible removal of the redacted information is achieved.

When identifying information that needs to be redacted prior to release, whole sentences or paragraphs should not be identified if only one or two words in that sentence or paragraph are to be redacted, unless the release would enable the identification of the redacted information by context.

Where necessary, information relating to the effect that a digital document has been redacted shall be linked with the digital document. To identify the fact that a redaction process has been undertaken, redacted information may be replaced by a sentence stating that some information has been redacted.

When redaction is performed on a digital document, any metadata included within the digital document shall be reviewed for redaction requirements and appropriate redactions undertaken.

Where redaction is performed on a digital document that contains images, video and/or voice information, redaction techniques that remove the necessary information shall be used.

5.2 Redaction principles

The redaction of digital documents shall be carried out in accordance with the following principles:

— Retention of digital original document

The original or master version of a digital document shall not be redacted – redaction shall be carried out on a copy of the digital document. Original digital documents (e.g. the un-redacted document) shall be retained and be accessible only to those authorized.

— Complete removal of redacted information

Redaction shall irreversibly remove the required information from the redacted version of the document. The information shall be completely removed from the digital document, not simply from the displayable content.

— Security evaluated redaction

Redaction shall always be carried out using methods approved by the organisation.

— Controlled environment

Electronic redaction shall be carried out in an environment that provides access only to those trained and authorised to carry out redaction.

— Intermediary stages

All intermediary stages of the redaction process should be deleted. Only the original digital document and the appropriately redacted version should be retained. Where a particular digital document is to be redacted in different ways (for example for different audiences), then it may be appropriate to temporarily retain intermediate stages until all redaction processes have been completed.

If there are many digital documents that may have many redactions for different reasons, and if over time some of the reasons for redaction expire, it may be necessary to create and retain over time an intermediary copy that indicates the text or objects that should be redacted and the reasons for the redaction. The information is actually redacted when the redacted copy is produced. This provides the reviewers with the capability to re-review the intermediary copy of redacted documents, and remove redactions whose reason for redaction has expired, without having to re-review the documents for unexpired reasons for redaction.

6 Redaction processes

6.1 Introduction

The redaction of digital documents is a relatively new area of information and records management practice, and raises unique issues and potential risks.