# ISO/IEC JTC 1 N 9962

2010-01-21

## ISO/IEC JTC 1
## Information Technology

**Document Type:** New Work Item Proposal

**Document Title:** Proposal for a new work item on Specification for digital redaction

**Document Source:** SC 27 Secretariat

**Reference:**

**Document Status:** This document is circulated to JTC 1 National Bodies for concurrent review. If the JTC 1 Secretariat receives no objections to this proposal by the due date indicated, we will so inform the SC 27 Secretariat

**Action ID:** Act

**Due Date:** 2010-04-21

**No. of Pages:** 17

REPLACES: N

**ISO/IEC JTC 1/SC 27**

**Information technology - Security techniques**

**Secretariat: DIN, Germany**

| | |
|---|---|
| **DOC TYPE:** | Proposed NP |
| **TITLE:** | **Proposal for a new work item on Specification for digital redaction** |
| **SOURCE:** | Secretariat of JTC 1/SC 27 |
| **DATE:** | 2010-01-18 |
| **PROJECT:** | **NWIP** |
| **STATUS:** | In accordance with resolution 2 (contained in SC 27 N7908) of the 7th SC 27/WG 4 meeting held in Redmond (WA, USA) 2nd - 6th November 2009, this document is being circulated to the SC 27 National Bodies for a 3-month NWIP letter ballot and to JTC 1 for a concurrent review.<br>P-Members of SC 27 are requested to submit their votes on this document via the ISO e-balloting application by **2010-04-18.** |
| **ACTION ID:** | LB |
| **DUE DATE:** | **2010-04-18** |
| **DISTRIBUTION:** | P- and L-Members<br>L. Rajchel, JTC 1 Secretariat<br>K. Brannon, ITTF<br>W. Fumy, SC 27 Chairman, M. De Soete, SC 27 Vice-Chair<br>E. J. Humphreys, K. Naemura, M. Bañón, M.-C. Kang, K. Rannenberg, WG-Conveners |
| **MEDIUM:** | Livelink-server |
| **NO. OF PAGES:** | 1 + 4 + 10 (Appendix = Outline document) |

Secretariat ISO/IEC JTC 1/SC 27 –
DIN Deutsches Institut für Normung e. V., Burggrafenstr. 6, 10772 Berlin, Germany
Telephone: + 49 30 2601-2652; Facsimile: + 49 30 2601-1723; E-mail: krystyna.passia@din.de;
HTTP://www.jtc1sc27.din.de/en

## New Work Item Proposal

## PROPOSAL FOR A NEW WORK ITEM

| | |
|---|---|
| Date of presentation of proposal:<br>20xx-xx-xx | Proposer: ISO/IEC JTC 1 SC 27 |
| Secretariat: ISO/IEC JTC 1/SC27<br>DIN, Germany | **ISO/IEC JTC 1/SC 27 N7939** |

**A proposal for a new work item** shall be submitted to the secretariat of the ISO/IEC joint technical committee concerned with a copy to the ISO Central Secretariat.

**Presentation of the proposal**

**Title: Information technology – Security techniques – Specification for Digital Redaction**

**Scope: Specification for Digital Redaction** - The redaction of born-digital records is a relatively new area of records management practice, and raises unique issues and potential risks. Records may be redacted electronically in their original format. This may be carried out either using deletion tools within the creating software, or by using specialised redaction software. This approach must be treated with extreme caution, due to the possibility that deleted information may still be recoverable, and the potential for information to remain hidden within non-displayable portions of the bit stream.

This standard would provide good practice guidelines for digital redaction as well as a testing methodology for evaluating the functionality of redacting functions of office software or separate dedicated redaction software solutions.

**Purpose and justification:**

When redacting electronically, great care must be taken over the choice of target format. It is crucial that no evidence of redacted information is retained in a redacted copy. Some binary formats may allow changes to be rolled back.

The simplest type of electronic record to redact is a plain text file, in which there is a one to one correspondence between bytes and displayable characters. Because of this direct correspondence, redacting these formats is simply a matter of deleting the displayed information - once the file is saved, the deleted information cannot be recovered.

However, the majority of electronic records created using modern office software are stored in proprietary, binary-encoded or XML based formats. Neither the XML nor the binary formats have the simple and direct correlation of plain text, and may contain significant information which is not displayed to the user, and the presence of which may therefore not be apparent. They may incorporate change histories, audit trails, or embedded metadata, by means of which deleted information can be recovered, or simple redaction processes otherwise circumvented. While the XML based formats are subject of international standardisation, the binary formats are usually the property of the software vendor which develops them. The mechanisms by which information is stored within these formats are often poorly understood. In addition, cryptographic and semantic analysis techniques can potentially be used to identify redacted information.

**Programme of work**

If the proposed new work item is approved, which of the following document(s) is (are) expected to be developed?
_x _ a single International Standard
____ more than one International Standard (expected number: ........ )
____ a multi-part International Standard consisting of .......... parts
____ an amendment or amendments to the following International Standard(s) ...................................
____ a technical report , type ...........

And which standard development track is recommended for the approved new work item?

_x_a. Default Timeframe

___b. Accelerated Timeframe

___c. Extended Timeframe

**Relevant documents to be considered:**

**Co-operation and liaison:**

**Preparatory work offered with target date(s)**

**WD 2010-05 CD 2011-05 FDIS 2012-10 IS 2013-05**

**Signature:** DIN, German NB of ISO/IEC JTC 1/SC 27

Will the service of a maintenance agency or registration authority be required: no
- If yes, have you identified a potential candidate?
- If yes, indicate name

Are there any known requirements for coding? no

-If yes, please specify on a separate page

Does the proposed standard concern known patented items? no
- If yes, please provide full information in an annex

Are there any known accessibility requirements and or dependencies (see: http://www.jtc1access.org)? ...........no.......
- If yes, please specify on a separate page

Are there any known requirements for cultural and linguistic adaptability? ........no............
- If yes, please specify on a separate page

**Comments and recommendations of the JTC 1 or SC27**- attach a separate page as an annex, if necessary

**Comments with respect to the proposal in general, and recommendations thereon:**
It is proposed to assign this new item to JTC 1/SC 27

**Voting on the proposal** - Each P-member of the ISO/IEC/JTC 1/SC 27 has an obligation to vote within the time limits laid down (normally three months after the date of circulation).

| | | |
|---|---|---|
| **Date of circulation:**<br>2010-01-18 | **Closing date for voting:**<br>2010-04-18 | **Signature of Secretary:**<br>Krystyna Passia<br>Secretariat JTC 1/SC27 |

| ***NEW WORK ITEM PROPOSAL - PROJECT ACCEPTANCE CRITERIA*** | | |
|---|---|---|
| **Criterion** | **Validity** | **Explanation** |
| **A. Business Requirement** | | |
| A.1 Market Requirement | Essential __<br>Desirable _x__<br>Supportive ___ | Generally accepted need for improved security in digital redaction. |
| A.2 Regulatory Context | Essential ___<br>Desirable _X__<br>Supportive __<br>Not Relevant __ | This standard might be used by evaluation facilities performing in support of regulatory requirements. |
| **B. Related Work** | | |
| B.1 Completion/Maintenance of current standards | Yes ___<br>No _x_ | |
| B.2 Commitment to other organization | Yes ___<br>No _x_ | |
| B.3 Other Source of standards | Yes ___<br>No _x_ | |
| **C. Technical Status** | | |
| C.1 Mature Technology | Yes _x_<br>No ___ | Implementations exist that perform digital redaction in various levels of sophistication |
| C.2 Prospective Technology | Yes __<br>No _x_ | |
| C.3 Models/Tools | Yes _x_<br>No __ | Testing methodology |
| **D. Conformity Assessment and Interoperability** | | |
| D.1 Conformity Assessment | Yes _x_<br>No __ | Testing methodology |
| D.2 Interoperability | Yes __<br>No _x_ | Guidelines and methodology are technology neutral |
| **E. Adaptability to Culture, Language, Human Functioning and Context of Use** | | |
| E1. Cultural and Linguistic Adaptability | Yes _x__<br>No __ | Testing methodology to take international character sets into account |

| | | |
|---|---|---|
| E.2 Adaptability to Human Functioning and Context of Use | Yes ___<br>No _x_ | |
| **F. Other Justification** | | |

Appendix to

ISO/IEC JTC 1/SC 27 N7939

# Title: Outline document for the New Work Item Proposal on Specification for digital redaction

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

## Introduction

Redaction is the separation of disclosable from non-disclosable information by blocking out individual words, sentences or paragraphs or the removal of whole pages or sections prior to the release of the document. In the paper environment some organisations will know redaction as extracts when whole pages are removed, or deletions where only a section of text is affected.

# 1 Scope

This standard provides best practices for performing redaction of electronic documents as well as methodology for testing the security of built-in in redaction functions of document preparation software or separate standalone electronic redaction tools.

**1.1** Purpose
The redaction of born-digital records is a relatively new area of records management practice, and raises unique issues and potential risks. Records may be redacted electronically in their original format. This may be carried out either using deletion tools within the creating software, or by using specialised redaction software. This approach must be treated with extreme caution, due to the possibility that deleted information may still be recoverable, and the potential for information to remain hidden within non-displayable portions of the bit stream.

**1.2** Target Audience
Personnel assigned with task of redacting documents as well as persons evaluating electronic document redaction built-in functions of office packages or standalone software tools.

# 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.

# 3 Terms and definitions

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

NOTE: This clause contains only those terms which are used in a specialised way throughout ISO/IEC xxxxx. The majority of terms in ISO/IEC xxxx are used either according to their accepted dictionary definitions or according to commonly accepted definitions that may be found in existing standards, ISO security glossaries or other well-known collections of security and privacy terms. Some combinations of common terms used in ISO/IEC xxxx, while not meriting inclusion in this clause x, are explained for clarity in the context where they are used.

**3.1** Document
definition

**3.2** Reviewer
definition

**3.3** Target
definition

**3.4** Source
definition

**3.5** Term
definition

## 4 Symbols and abbreviated Terms

The following abbreviations are common to more than one part of ISO/IEC xxxxx

## 5 Principles of redaction

Redaction should always be reversible - it should never result in permanent removal of text. Redaction should always be carried out on copies, whether paper or electronic.

Redaction is carried out in order to edit exempt details from a document. It should be used when one or two individual words, a sentence or paragraph, a name, address or signature needs to be removed.

If so much information has to be withheld that a document becomes nonsensical, the entire document should be withheld. In the case of paper documents the same principle should apply to individual pages.

When undertaking redaction, reviewers should consider whether any other factors are important for the understanding of the material. For example, if colour makes meaning clear in a paper document, a redacted colour copy should be released.

Redaction should be performed or overseen by staff that are knowledgeable about the records and can determine what material is exempt. If those staff identifying such material do not carry out redaction themselves, their instructions must be specific e.g. 'Memo dated …, paragraph no…, line starting… and ending…' etc.

## 6 Identifying material for redaction

All organisations should have staff able to identify information that may be exempt from being redacted. Ideally they should have a good knowledge of the records being reviewed for release.

In order to conform fully with requests for information, it is essential that only exempt material be redacted. A whole sentence or paragraph should not be removed if only one or two words are non-disclosable, unless release would place the missing words in context and make their content or meaning clear.

Reviewers should also consider that earlier statements in a document might suggest the content of removed material. For example, if a paragraph refers to reports from overt sources, and the following paragraph refers to reports from covert sources, as well as removing the words 'covert sources', 'overt sources' would also need to be removed or the meaning of the missing words from the second paragraph could be inferred.

## 7 Keeping records of redaction work

Retained electronic (unredacted) documents must be kept in a secure area of the electronic file plan or local area network. They should be accessible only to authorised staff.

## 8 Tracking of retained redactions

Organisations performing redaction should maintain a database of documents that have been redacted. The database should include the following fields.

## 9 Redaction of electronic records

### **9.1** Introduction

The following discusses the technical aspects of redacting electronic records. It should be remembered that when dealing with electronic records the general principles of redaction are the same as those described in “Principles of Redaction” in this document.

### **9.2** Issues in redacting electronic records

The redaction of born-digital records is a relatively new area of records management practice, and raises unique issues and potential risks. Records may be redacted electronically in their original format. This may be carried out either using deletion tools within the creating software, or by using specialised redaction software. This approach must be treated with extreme caution, due to the possibility that deleted information may still be recoverable, and the potential for information to remain hidden within non-displayable portions of the bit stream.

### **9.3** Simple electronic redaction

The simplest type of electronic record to redact is a plain text file, in which there is a one to one correspondence between bytes and displayable characters. Because of this direct correspondence, redacting these formats is simply a matter of deleting the displayed information - once the file is saved, the deleted information cannot be recovered.

#### **9.3.1** Character encoding

Care must be taken with regards to encoding of characters used in the plain text file. Extended encoding mechanisms such as Unicode require appropriate editor is used, otherwise the one to one correspondence between bytes and displayable characters is lost.

### **9.4** Complex electronic redaction

When redacting electronically, great care must be taken over the choice of target format. It is crucial that no evidence of redacted information is retained in a redacted copy. Some binary formats may allow changes to be rolled back.

The majority of electronic records created using modern office software are stored in proprietary, binary-encoded or XML based formats. Neither the XML nor the binary formats have the simple and direct correlation of plain text, and may contain significant information which is not displayed to the user, and the presence of which may therefore not be apparent. They may incorporate change histories, audit trails, or embedded metadata, by means of which deleted information can be recovered, or simple redaction processes otherwise circumvented. While the XML based formats are subject of international standardisation, the binary formats are usually the property of the software vendor which develops them. The mechanisms by which information is stored within these formats are often poorly understood. In addition, cryptographic and semantic analysis techniques can potentially be used to identify redacted information.

## 10 Principles to electronic redaction

The redaction of electronic records should always be carried out in accordance with the following principles:

### 10.1 Retention of original

The original or master version of an electronic record must never be redacted – redaction must always be carried out on a new copy of the record, either in paper or electronic format.

### 10.2 Complete removal of redacted information

Redaction must irreversibly remove the required information from the redacted copy of the record. The information must be completely removed from the bit stream, not simply from the displayable record.

### 10.3 Security evaluated redaction

Redaction should always be carried out using methods which have been fully security tested.

### 10.4 Controlled environment

Electronic redaction should be carried out in a controlled and secure environment that provides access only to those trained and authorised to carry out redaction.

### 10.5 Intermediary stages

All intermediary stages of the redaction process should be deleted. Only the original record and the appropriately redacted copy should be retained.

## 11 Approaches to electronic redaction

A number of different approaches to electronic redaction are possible.

### 11.1 Traditional redaction

For electronic records, which can be printed as a hardcopy, traditional redaction techniques, can be applied. Either the record may be printed and redaction carried out on the printed copy, or the information may be redacted from an electronic copy, which is then printed. If the redacted copy is required in electronic format, this can be created by scanning the redacted paper copy into an appropriate format.

### 11.2 Format redaction

Records may be redacted electronically in their original format. This may be carried out either using deletion tools within the creating software, or by using specialised redaction software. This approach must be treated with extreme caution, due to the possibility that deleted information may still be recoverable, and the potential for information to remain hidden within non-displayable portions of the bit stream.

### 11.3 Conversion

An electronic record may be redacted through a combination of information deletion and conversion to a different format. Certain formats, such as plain ASCII text files, contain displayable information only. Conversion to this format will therefore eliminate any information that may be hidden in non-displayable portions of a bit stream.

### 11.4 Roundtrip redaction

The redacted record may be required to be made available in its original format, for example, to preserve complex formatting. In such cases, an extension of the conversion approach may be applicable. Roundtripping entails the conversion of the record to another format, followed by conversion back to the original format, such that the conversion process removes all evidence of the redacted information. Information deletion may be carried out either prior to conversion, or in the intermediary format.

This approach requires a thorough understanding of the formats and conversion processes involved, and the mechanisms by which information is transferred during conversion.

#### 11.4.1 Electronically redacting documents

When redacting electronically, great care must be taken over the choice of target format. It is crucial that no evidence of redacted information is retained in a redacted copy. Some binary formats may allow changes to be rolled back; consequently these formats should not be used for creating redacted copies.

PDF can be used as a format for redacted copies, but PDF files should be roundtripped via a simple image format to ensure that all evidence of previously redacted information is removed. A simple lossless image format, such as Windows Bitmap (BMP), should be used, as it contains no provision for storing metadata. There is therefore no means by which hidden information could be inserted into the image file. This format has been preferred over other image formats such as TIFF for this reason. The TIFF format contains metadata not visible on screen.

### 11.5 Redacting office documents

This guidance applies to word-processed documents, such as documents created using all versions of Microsoft Word, WordPerfect and OpenOffice Writer, and to spreadsheets, such as those created using all versions of Microsoft Excel, Lotus 1-2-3, and OpenOffice Calc.

### 11.6 Redacting PDF documents

## 12 Mechanisms for testing the security of electronic redaction techniques

## Notes to Proforma

**A. Business Relevance.** That which identifies market place relevance in terms of what problem is being solved and or need being addressed.

A.1 Market Requirement. When submitting a NP, the proposer shall identify the nature of the Market Requirement, assessing the extent to which it is essential, desirable or merely supportive of some other project.

A.2 Technical Regulation. If a Regulatory requirement is deemed to exist - e.g. for an area of public concern e.g. Information Security, Data protection, potentially leading to regulatory/public interest action based on the use of this voluntary international standard - the proposer shall identify this here.

**B. Related Work.** Aspects of the relationship of this NP to other areas of standardisation work shall be identified in this section.

B.1 Competition/Maintenance. If this NP is concerned with completing or maintaining existing standards, those concerned shall be identified here.

B.2 External Commitment. Groups, bodies, or for a external to JTC 1 to which a commitment has been made by JTC for Co-operation and or collaboration on this NP shall be identified here.

B.3 External Std/Specification. If other activities creating standards or specifications in this topic area are known to exist or be planned, and which might be available to JTC 1 as PAS, they shall be identified here.

**C. Technical Status.** The proposer shall indicate here an assessment of the extent to which the proposed standard is supported by current technology.

C.1 Mature Technology. Indicate here the extent to which the technology is reasonably stable and ripe for standardisation.

C.2 Prospective Technology. If the NP is anticipatory in nature based on expected or forecasted need, this shall be indicated here.

C.3 Models/Tools. If the NP relates to the creation of supportive reference models or tools, this shall be indicated here.

### D. Conformity Assessment and Interoperability

D.1 Indicate here if Conformity Assessment is relevant to your project. If so, indicate how it is addressed in your project plan.

D.2 Indicate here if Interoperability is relevant to your project. If so, indicate how it is addressed in your project plan

### E. Adaptability to Culture, Language, Human Functioning and Context of Use

NOTE: The following criteria do not mandate any feature for adaptability to culture, language, human functioning or context of use. The following criteria require that if any features are provided for adapting to culture, language, human functioning or context of use by the new Work Item proposal, then the proposer is required to identify these features.

E.1 Cultural and Linguistic Adaptability. Indicate here if cultural and natural language adaptability is applicable to your project. If so, indicate how it is addressed in your project plan. ISO/IEC TR 19764 (Guidelines, methodology, and reference criteria for cultural and linguistic adaptability in information technology products) now defines it in a simplified way:

“ability for a product, while keeping its portability and interoperability properties, to:

- be internationalized, that is, be adapted to the special characteristics of natural languages and the commonly accepted rules for their se, or of cultures in a given geographical region;

- take into account the usual needs of any category of users, with the exception of specific needs related to physical constraints”

*Examples of characteristics of natural languages are: national characters and associated elements (such as hyphens, dashes, and punctuation marks), writing systems, correct transformation of characters, dates and measures, sorting and searching rules, coding of national entities (such as country and currency codes), presentation of telephone numbers and keyboard layouts. Related terms are localization, jurisdiction and multilingualism.*

E.2 Adaptability to Human Functioning and Context of Use. Indicate here whether the proposed standard takes into account diverse human functioning and diverse contexts of use. If so, indicate how it is addressed in your project plan.

NOTE:

1. Human functioning is defined by the World Health Organization at **http://www3.who.int/icf/beginners/bg.pdf as:**
   <<In ICF (International Classification of Functioning, Disability and Health), the term functioning refers to all body functions, activities and participation.>>

2. 2. Content of use is defined in ISO 9241-11:1998 (Ergonomic requirements for office work with visual display terminals (VDTs) – Part 11: Guidance on usability) as:
   <<Users, tasks, equipment (hardware, software and materials), and the physical and societal environments in which a product is used.>>

3. 3. Guidance for Standard Developers to address the needs of older persons and persons with disabilities).

**F. Other Justification** Any other aspects of background information justifying this NP shall be indicated here