FINAL DRAFT
Technical Specification

ISO/DTS 32004

Document management — Portable Document Format — Integrity protection in encrypted documents in PDF 2.0

Gestion des documents — Format de document portable — Protection de l'intégrité dans les documents chiffrés en PDF 2.0

ISO/TC **171**/SC **2**

Secretariat: **ANSI**

Voting begins on:
2024-02-02

Voting terminates on:
2024-03-29

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.

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.

Reference number
ISO/DTS 32004:2024(en)

ISO copyright office
CP 401 • Ch. de Blandonnet 8
CH-1214 Vernier, Geneva
Phone: +41 22 749 01 11
Email: copyright@iso.org
Website: www.iso.org

Published in Switzerland

Contents

Page

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.

The procedures used to develop this document and those intended for its further maintenance are described in the ISO/IEC Directives, Part 1. In particular, the different approval criteria needed for the different types of ISO document should be noted. This document was drafted in accordance with the editorial rules of the ISO/IEC Directives, Part 2 (see www.iso.org/directives).

ISO draws attention to the possibility that the implementation of this document may involve the use of (a) patent(s). ISO takes no position concerning the evidence, validity or applicability of any claimed patent rights in respect thereof. As of the date of publication of this document, ISO had not received notice of (a) patent(s) which may be required to implement this document. However, implementers are cautioned that this may not represent the latest information, which may be obtained from the patent database available at www.iso.org/patents. ISO shall not be held responsible for identifying any or all such patent rights.

Any trade name used in this document is information given for the convenience of users and does not constitute an endorsement.

For an explanation of the voluntary nature of standards, the meaning of ISO specific terms and expressions related to conformity assessment, as well as information about ISO's adherence to the World Trade Organization (WTO) principles in the Technical Barriers to Trade (TBT), see www.iso.org/iso/foreword.html.

This document was prepared by Technical Committee ISO/TC 171, *Document management applications*, Subcommittee SC 2, *Document file formats, EDMS systems and authenticity of information*.

Any feedback or questions on this document should be directed to the user’s national standards body. A complete listing of these bodies can be found at www.iso.org/members.html.

Introduction

When encrypting documents, it is important to not only preserve the confidentiality of the encrypted material, but also to ensure that the receiving party can verify its integrity. Encryption mechanisms defined in ISO 32000-2:2020 currently only provide confidentiality without this authentication aspect.

This document describes a mechanism to protect the integrity of an encrypted PDF document using a Message Authentication Code (MAC), with key material derived from the file encryption key. Message authentication codes are distinct from digital signatures based on public-key cryptography. Digital signatures and message authentication codes have different but complementary security properties: a valid MAC created following this document proves knowledge of the file encryption key, whereas digital signatures as defined in ISO 32000-2:2020 do not have that property.

The MAC mechanism described in this document is backwards compatible with ISO 32000-2:2020 and can also be used in PDF documents containing digital signatures.

This document follows the lexical conventions regarding the usage of bold and italics which are specified in ISO 32000-2:2020, Clause 4.

Document management — Portable Document Format — Integrity protection in encrypted documents in PDF 2.0

1 Scope

This document specifies how to extend the ISO 32000-2:2020 specification by adding extensions to the **Encrypt** dictionary and trailer dictionary to provide integrity protection to the encrypted PDF document. This document also ensures that extensions are fully backward-compatible.

These extensions are intended for developers of software that creates PDF files (PDF writers), software that reads existing PDF files and (usually) interprets their contents for display (PDF readers), software that reads and displays PDF content and interacts with the computer users to possibly modify and save the PDF file (interactive PDF processors) and PDF products that read and/or write PDF files for a variety of other purposes (PDF processors).

NOTE PDF writers and PDF readers are more specialized classifications of interactive PDF processors and both are PDF processors.

This document does not specify the following:

- specific processes for converting paper or electronic documents to the PDF file format;
- specific technical design, user interface implementation, or operational details of rendering;
- specific physical methods of storing these documents such as media and storage conditions;
- methods for validating the conformance of PDF files or PDF processors;
- required computer hardware and/or operating system.

2 Normative references

The following documents are referred to in the text in such a way that some or all of their content constitutes requirements 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 32000-2:2020, *Document management — Portable document format — Part 2: PDF 2.0*

IETF RFC 2104, *HMAC: Keyed-Hashing for Message Authentication. [online]. 1997. Available from:*[1]

IETF RFC 4231, *Identifiers and Test Vectors for HMAC-SHA-224, HMAC-SHA-256, HMAC-SHA-384, and HMAC-SHA-512. [online]. 2005*[2]

IETF RFC 3394, *Advanced Encryption Standard (AES) Key Wrap Algorithm. [online]*[3]

IETF RFC 5652:2009, *Cryptographic Message Syntax (CMS). [online]. 2009*[4]

IETF RFC 5869, *HMAC-based Extract-and-Expand Key Derivation Function (HKDF). [online]. 2010*[5]

[1] https://tools.ietf.org/html/rfc2104.html

[2] https://tools.ietf.org/html/rfc4231.html

[3] https://tools.ietf.org/html/rfc3394.html

[4] https://tools.ietf.org/html/rfc5652.html

[5] https://tools.ietf.org/html/rfc5869.html

IETF RFC 6211, *Cryptographic Message Syntax (CMS): Algorithm Identifier Protection Attribute. [online]. 2011*[6]

NIST Computer Security Objects Register (CSOR). [online]. 2009. Available from:[7]

3 Terms and Definitions

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

ISO and IEC maintain terminology databases for use in standardization at the following addresses:

- ISO Online browsing platform: available at https://www.iso.org/obp
- IEC Electropedia: available at https://www.electropedia.org/

3.1
PDF
Portable Document Format
file format defined by ISO 32000-2:2020

3.2
MAC
Message Authentication Code
cryptographic check sum on data that uses a symmetric key to detect both accidental and intentional modification of data

[SOURCE: ISO 16609:2022, 3.10]

3.3
file encryption key
key used for document-level encryption by a PDF security handler of version 5 or higher

Note 1 to entry: General provisions about PDF security handlers can be found in ISO 32000-2:2020, 7.6.3.

Note 2 to entry: Extensions to PDF can define security handlers other than those specified in ISO 32000-2:2020. File encryption keys defined by such security handlers are also covered by this definition.

3.4
signed revision
initial or incremental revision of a *PDF* (3.1) file that adds a digital signature or document time stamp signature to the document

3.5
unsigned revision
initial or incremental revision of a *PDF* (3.1) file that does not add a digital signature or document time stamp signature to the document

Note 1 to entry: A PDF document can contain both signed and unsigned revisions.

3.6
Abstract Syntax Notation One
ASN.1
International Standard for representing data types and structures

Note 1 to entry: The encoding rules for this abstract syntax notation are defined in ISO/IEC 8825-1.

[SOURCE: ISO 17261:2012, 3.5, modified — ISO/IEC 8825-2 has been replaced by ISO/IEC 8825-1 in the Note 1 to entry.]

[6] https://tools.ietf.org/html/rfc6211.html

[7] https://csrc.nist.gov/projects/computer-security-objects-register

3.7
distinguished encoding rules
DER
encoding rules that may be applied to values of types defined using the *ASN.1* (3.6) notation

[SOURCE: ISO/IEC 18014-2:2021, 3.22, modified — note 1 to entry has been removed]

3.8
hashed message authentication code
HMAC
mechanism for message authentication using a cryptographic hash function in combination with a shared secret key

Note 1 to entry: This definition has been adapted from RFC 2104.

4 Extension schema details

The developer extensions dictionary in Table 1 shall be part of the document's extensions dictionary (ISO 32000-2:2020, 7.12). It shall be included as an array entry under the **ISO_** prefix.

Encrypted PDF documents making use of the extension specified in this document shall conform to ISO 32000-2:2020.

Table 1 — Extension schema entries

Key	**Type**	**Value**
Type	name	*DeveloperExtensions*
BaseVersion	name	*2.0*
ExtensionLevel	integer	32004
ExtensionRevision	text string	:2024 NOTE The COLON (U+003A) character is part of the revision identifier.
URL	string	https://www.iso.org/standard/45877.html

5 Proposed changes

5.1 Encrypt dictionary

5.1.1 Additions to ISO 32000-2:2020, 7.6.2

The content of Table 2 is appended to ISO 32000-2:2020, Table 20.

Table 2 — Additions to ISO 32000-2:2020, Table 20

Key	**Type**	**Value**
KDFSalt	byte string	*(Conditionally required; shall be a direct object)* A 32-byte salt value for use in key derivation (see 6.4). This entry is required in documents that make use of PDF MAC.

NOTE The value of the **KDFSalt** entry is intended to remain constant throughout all incremental updates of the document.

5.1.2 Additions to ISO 32000-2:2020, 7.6.4.2

The content of Table 3 is appended to ISO 32000-2:2020, Table 22.

Table 3 — Additions to ISO 32000-2:2020, Table 22

Bit position	Meaning
13	When zero, indicates that a PDF MAC token is required to be present in all revisions of the document. The location of the PDF MAC token is indicated by the **AuthCode** dictionary (see 5.2.3).

NOTE 1 This addition supersedes the provision of ISO 32000-2:2020, 7.6.4.2 requiring that PDF readers ignore all flags other than those at bit positions 3, 4, 5, 6, 9, 10, 11 and 12.

NOTE 2 The intention behind this permission bit is to signal to a PDF processor reading the document that a PDF MAC token is expected. The encrypted **Perms** entry provides a degree of tamper-resistance and helps to protect the document against attackers stripping the MAC. This protection is not without limitations: unless bit 13 is zero in *all* revisions, it can be trivially bypassed by a knowledgeable adversary.

5.1.3 Additions to ISO 32000-2:2020, 7.6.5.2

The content of Table 4 is appended to ISO 32000-2:2020, Table 24.

Table 4 — Additions to ISO 32000-2:2020, Table 24

Bit position	Meaning
13	When zero, indicates that a PDF MAC token is required to be present in all revisions of the document. The location of the PDF MAC token is indicated by the **AuthCode** dictionary (see 5.2.3).

NOTE This addition supersedes the provision of ISO 32000-2:2020, 7.6.5.2 requiring that PDF readers ignore all flags other than those at bit positions 2, 3, 4, 5, 6, 9, 10, 11 and 12.

5.2 File trailer

5.2.1 Additions to ISO 32000-2:2020, 7.5.5

The content of Table 5 is appended to ISO 32000-2:2020, Table 15.

Table 5 — Additions to ISO 32000-2:2020, Table 15

Key	Type	Value
AuthCode	dictionary	*(Required if the document is encrypted with user access permissions bit 13 zero. Shall be a direct object; PDF 2.0)* Describes a PDF MAC token to validate the integrity of an encrypted document (see 5.2.3). If present, the value of the **V** entry in the document's **Encrypt** dictionary shall be at least 5.

5.2.2 Additions to ISO 32000-2:2020, 7.6.2

The following entry is added to the bulleted list in ISO 32000-2:2020, 7.6.2.

— Any byte strings representing the value of the **MAC** key in an **AuthCode** dictionary.

NOTE This prevents strings containing PDF MAC tokens from being encrypted.

5.2.3 AuthCode dictionary

The **AuthCode** dictionary, defined in Table 6, contains a PDF MAC token or describes where to find it. All **AuthCode** dictionary entries defined below shall be direct objects, with the exception of the **SigObjRef** entry.