



**Document management—Electronic
document file format for long-term
preservation**

**Part 3: Use of ISO 32000-1 with support
for embedded files (PDF/A-3)**



This Australian Standard® was prepared by Committee IT-021, Records and Document Management Systems. It was approved on behalf of the Council of Standards Australia on 20 March 2017.

This Standard was published on 1 May 2017.

The following are represented on Committee IT-021:

- AMS Imaging
 - Australian Computer Society
 - Australian Industry Group
 - Australian Society of Archivists
 - City of Boroondara
 - Department of Defence (Australian Government)
 - Monash University
 - National Archives of Australia
 - Public Record Office Victoria
 - Queensland State Archives
 - Records and Information Management Professionals Australasia
 - State Archives and Records Authority of New South Wales
 - State Records Office of Western Australia
 - University of Sydney
-

This Standard was issued in draft form for comment as DR AS ISO 19005.3:2017.

Standards Australia wishes to acknowledge the participation of the expert individuals that contributed to the development of this Standard through their representation on the Committee and through the public comment period.

Keeping Standards up-to-date

Australian Standards® are living documents that reflect progress in science, technology and systems. To maintain their currency, all Standards are periodically reviewed, and new editions are published. Between editions, amendments may be issued.

Standards may also be withdrawn. It is important that readers assure themselves they are using a current Standard, which should include any amendments that may have been published since the Standard was published.

Detailed information about Australian Standards, drafts, amendments and new projects can be found by visiting www.standards.org.au

Standards Australia welcomes suggestions for improvements, and encourages readers to notify us immediately of any apparent inaccuracies or ambiguities. Contact us via email at mail@standards.org.au, or write to Standards Australia, GPO Box 476, Sydney, NSW 2001.

Australian Standard®

**Document management—Electronic
document file format for long-term
preservation**

**Part 3: Use of ISO 32000-1 with support
for embedded files (PDF/A-3)**

First published as AS ISO 19005.3:2017.

COPYRIGHT

© Standards Australia Limited

All rights are reserved. No part of this work may be reproduced or copied in any form or by any means, electronic or mechanical, including photocopying, without the written permission of the publisher, unless otherwise permitted under the Copyright Act 1968.

Published by SAI Global Limited under licence from Standards Australia Limited, GPO Box 476, Sydney, NSW 2001, Australia

ISBN 978 1 76035 729 0

PREFACE

This Standard was prepared by the Standards Australia Committee IT-021, Records and Document Management Systems.

The objective of this Standard is to specify the use of the portable document format (PDF) 1.7, as formalized in ISO 32000-1, for preserving the static visual representation of page-based electronic documents over time in addition to allowing any type of other content to be included as an embedded file or attachment.

This Standard is identical with, and has been reproduced from ISO 19005-3:2012, *Document management — Electronic document file format for long-term preservation — Part 3: Use of ISO 32000-1 (PDF/A-2)*.

As this Standard is reproduced from an International Standard, the following applies:

- (a) In the source text 'this part of ISO 19005 should read 'this Australian Standard'.
- (b) A full point substitutes for a comma when referring to a decimal marker.

None of the normative references in the source document have been adopted as Australian or Australian/New Zealand Standards.

The terms 'normative' and 'informative' have been used in this Standard to define the application of the annex to which they apply. A 'normative' annex is an integral part of a Standard, whereas an 'informative' annex is only for information and guidance.

CONTENTS

1	Scope	1
2	Normative references	1
3	Terms and definitions	2
4	Notation	3
5	Conformance levels	4
	5.1 General.....	4
	5.2 Level A conformance.....	4
	5.3 Level B conformance.....	5
	5.4 Level U conformance.....	5
	5.5 Conforming readers.....	5
6	Technical requirements	6
	6.1 File structure.....	6
	6.2 Graphics.....	8
	6.3 Annotations.....	17
	6.4 Interactive forms.....	18
	6.5 Action.....	19
	6.6 Metadata.....	19
	6.7 Logical structure.....	24
	6.8 Embedded files.....	26
	6.9 Optional content.....	27
	6.10 Use of alternate presentations and transitions.....	28
	6.11 Document requirements.....	28
	Annex A (normative) Method for determining transparency on a page	29
	Annex B (normative) Requirements for digital signatures in PDF/A	31
	Annex C (informative) Best practices for PDF/A	32
	Annex D (informative) Incorporation of XFA data sets	34
	Annex E (normative) Associated files	35
	Bibliography	41

INTRODUCTION

PDF is a digital format for representing page-based documents. PDF files can be created natively in PDF form, converted from other electronic formats or digitized from paper, microform, or other hard copy format. Businesses, governments, libraries, archives and other institutions and individuals around the world use PDF to represent considerable bodies of important information. Much of this information needs to be kept for substantial lengths of time; some needs to be kept permanently. These PDF files need to remain useable and accessible across multiple generations of technology. However, the inclusive, feature-rich nature of the format requires that constraints be placed on its use to make it suitable for the long-term preservation of electronic documents. The future use of, and access to, these objects depends upon maintaining their visual appearance as well as their higher-order properties, such as the logical organization of pages, sections, and paragraphs, machine recoverable text stream in natural reading order, and a variety of administrative, preservation and descriptive metadata.

This International Standard is created as a multi-part document, of which this is part 3. This allows future parts to be created without rendering this document or applications based on this document obsolete.

The primary purpose of ISO 19005 is to define a file format based on PDF, known as PDF/A, which provides a mechanism for representing electronic documents in a manner that preserves their static visual appearance over time, independent of the tools and systems used for creating, storing or rendering the files.

A secondary purpose of ISO 19005 is to define a framework for representing the logical structure and other semantic information of electronic documents within conforming files.

Another purpose of ISO 19005 is to provide a framework for recording the context and history of electronic documents in metadata within conforming files.

This part of ISO 19005 adds a new goal (beyond that of ISO 19005-2) which is to enable PDF documents to serve as containers for other file formats, so that a single physical file can contain not only the visual representation but also other representations including the original authored version, richer semantic formats, and others. This part of ISO 19005 does not address the long-term suitability of formats, that may be embedded, other than those compliant with any part of this International Standard.

These goals are accomplished by identifying the set of PDF components that can be used, and restrictions on the form of their use, within conforming PDF/A files.

By itself, PDF/A does not necessarily ensure that the visual appearance of the content accurately reflects any original source material used to create the conforming file, e.g. the process used to create a conforming file might substitute fonts, reflow text, downsample images or use lossy compression. Organizations that need to ensure that a conforming file is an accurate representation of original source material might need to impose additional requirements on the processes that generate the conforming file beyond those imposed by this part of ISO 19005, such as those best practices in Annex C. In addition, it is important for those organizations to implement policies and practices regarding the inspection of conforming files for correct visual appearance.

PDF/A does not directly address the topic of authenticity either for the underlying content to be visually represented or the PDF/A file itself. Such authenticity is generally considered to be important for legal, regulatory and governance purposes and is beyond the scope of this part of ISO 19005.

This part of ISO 19005 is one component of an organization's electronic archival environment for long-term retention of documents. Successful implementation of this part of ISO 19005 for archival purposes depends upon:

- the retention requirements of an organization's archival environment, records management policies and procedures as specified in ISO 15489-1^[6];
- any additional conditions necessary to ensure the persistence of electronic documents and their characteristics over time, including, but not limited to, those defined by ISO 14721^[5], ISO/TR 15801^[7], and ISO/TR 18492^[8];

- quality assurance processes necessary to verify conformance with applicable requirements and conditions, e.g. an inspection regime to verify the quality and integrity of converted source data.

This part of ISO 19005 is intended to lead to the development of various applications that read, render, write and validate conforming files. Different applications will incorporate various capabilities to prepare, interpret and process conforming files based on needs as perceived by the suppliers of those applications. However, it is important to note that a conforming application needs to be able to read and process appropriately all files complying with a specified conformance level.

This part of ISO 19005 extends the capabilities of ISO 19005-2. Just as with ISO 19005-2, it is based on PDF version 1.7 (as defined in ISO 32000-1).

This part of ISO 19005 (in conjunction with its normative references) provides sufficient information to interpret any conforming PDF/A-3 file.

NPES and AIIM (accredited standards developing organizations) maintain an ongoing series of application notes for guiding developers and users of this part of ISO 19005. These application notes are available at <http://www.npes.org/standards/toolspdfa.html> and <http://www.aiim.org/Research-and-Publications/Standards/Articles/PDFA-Application-Notes>. Both NPES and AIIM will also retain copies of the specific non-ISO normative references of this part of ISO 19005 which are publicly available electronic documents.

AUSTRALIAN STANDARD

Document management—Electronic document file format for long-term preservation

Part 3:

Use of ISO 32000-1 with support for embedded files (PDF/A-3)

1 Scope

This part of ISO 19005 specifies the use of the Portable Document Format (PDF) 1.7, as formalized in ISO 32000-1, for preserving the static visual representation of page-based electronic documents over time in addition to allowing any type of other content to be included as an embedded file or attachment.

This part of ISO 19005 is not applicable to:

- specific processes for converting paper or electronic documents to the PDF/A 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;
- required computer hardware and/or operating systems.

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 646, *Information technology — ISO 7-bit coded character set for information interchange*¹⁾

ISO/IEC 10646, *Information technology — Universal Coded Character Set (UCS)*²⁾

ISO 15076-1, *Image technology colour management — Architecture, profile format and data structure — Part 1: Based on ICC.1:2010*

ISO 15930-7:2010, *Graphic technology — Prepress digital data exchange using PDF — Part 7: Complete exchange of printing data (PDF/X-4) and partial exchange of printing data with external profile reference (PDF/X-4p) using PDF 1.6*

ISO 19005-1, *Document management — Electronic document file format for long-term preservation — Part 1: Use of PDF 1.4 (PDF/A-1)*

ISO 19005-2, *Document management — Electronic document file format for long-term preservation — Part 2: Use of ISO 32000-1 (PDF/A-2)*

ISO 32000-1, *Document management — Portable document format — Part 1: PDF 1.7*

Extensible Markup Language (XML) 1.0 (Third Edition), W3C Recommendation, 4 February 2004. Available at <<http://www.w3.org/TR/2004/REC-xml-20040204>>

1) The character encoding defined in ISO/IEC 646 is equivalent to ANSI X3.4 (ASCII) and ECMA-6.

2) The character code values defined in ISO/IEC 10646 are equivalent to those of Unicode.