



中华人民共和国国家标准

GB/T 23286.3—2021/ISO 19005-3:2012

文献管理 长期保存的电子文档文件格式 第 3 部分：支持嵌入式文件的 ISO 32000-1 的使用(PDF/A-3)

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)

(ISO 19005-3:2012, IDT)

2021-04-30 发布

2021-11-01 实施

国家市场监督管理总局 发布
国家标准化管理委员会

目 次

前言	I
引言	II
1 范围	1
2 规范性引用文件	1
3 术语和定义	2
4 表示法	4
5 一致性级别	4
5.1 概述	4
5.2 A级一致性	5
5.3 B级一致性	5
5.4 U级一致性	5
5.5 合格阅读器	5
6 技术要求	5
6.1 文件结构	5
6.2 图形	8
6.3 标注	15
6.4 交互式表单	16
6.5 动作	16
6.6 元数据	17
6.7 逻辑结构	21
6.8 嵌入式文件	23
6.9 可选内容	24
6.10 使用替代演示和过渡	24
6.11 文档要求	24
附录 A (规范性附录) 页面上透明度的确定方法	25
附录 B (规范性附录) PDF/A 中数字签名要求	27
附录 C (资料性附录) PDF/A 的最佳实践	28
附录 D (资料性附录) XFA 数据集的编入	30
附录 E (规范性附录) 关联文件	31
附录 NA (资料性附录) 中英文词汇对照	38
参考文献	42

前 言

GB/T 23286《文献管理 长期保存的电子文档文件格式》拟分为若干部分：

- 第 1 部分：PDF1.4(PDF/A-1)的使用；
- 第 2 部分：ISO 32000-1 的使用(PDF/A-2)；
- 第 3 部分：支持嵌入式文件的 ISO 32000-1 的使用(PDF/A-3)。

本部分为 GB/T 23286 第 3 部分。

本部分按照 GB/T 1.1—2009 给出的规则起草。

本部分使用翻译法等同采用 ISO 19005-3:2012《文献管理 长期保存的电子文档文件格式 第 3 部分：支持嵌入式文件的 ISO 32000-1 的使用(PDF/A-3)》。

与本部分中规范性引用的国际文件有一致性对应关系的我国文件如下：

- GB/T 1988—1998 信息技术 信息交换用七位编码字符集(eqv ISO/IEC 646:1991)
- GB/T 13000—2010 信息技术 通用多八位编码字符集(UCS)(ISO/IEC 10646:2003, IDT)
- GB/T 23286.1—2009 文献管理 长期保存的电子文档文件格式 第 1 部分：PDF1.4(PDF/A-1)的使用(ISO 19005-1:2005, IDT)
- GB/T 32010.1—2015 文献管理 可移植文档格式 第 1 部分：PDF1.7(ISO 32000-1:2008, IDT)

本部分做了下列编辑性修改：

- 增加了资料性附录 NA。

请注意本文件的某些内容可能涉及专利。本文件的发布机构不承担识别这些专利的责任。

本部分由全国文献影像技术标准化技术委员会(SAC/TC 86)归口。

本部分起草单位：国家图书馆、福建福昕软件开发股份有限公司、北京影研创新科技发展有限公司。

本部分主要起草人：冯辉、曹宁、李志尧、刘芳铭、薛四新、龙伟、熊雨前、王坤、徐南。

引 言

可移植文档格式(Portable Document Format, PDF)是一种表现基于页面的文档数字格式。PDF文件可以原生创建,也可以从其他电子格式转换而来,或者从纸张、缩微印刷品或其他硬拷贝格式经过数字化处理而再生成。世界各地的企业、政府、图书馆、档案馆等机构和个人形成的大量 PDF 文件都有长久保存的需求,为了确保这些文件可以在不断的技术更新中持久保持可被有效查询和利用的功能,需要我们对其进行规范约束,使其具备必要的特殊属性,这其中既包括文件视觉外观的保持,也包括诸如页面、章节和段落的逻辑组织、按自然阅读顺序阅读的机器可恢复文本流的保持,以及各种管理性、保存性和描述性的元数据等的完备存现和保持。

GB/T 23286 系列标准等同采用 ISO 19005 的相应部分,其制定目的在于:

- 定义一种基于 PDF 的文件格式,称作 PDF/A。该格式提供了一个机制,用于长期保持电子文档静态视觉外观,而与创建、储存或呈现文件的工具和系统无关。
- 定义一个在合格文件中表现电子文档逻辑结构和其他语义信息的框架。
- 提供一个在合格文件中以元数据记录电子文档背景和历史的框架。

本部分增加了一个新目标(在第 2 部分的目标之外),即允许 PDF 文档充当其他文件格式的容器,使得单个物理文件不仅可以保持视觉表现始终如一,还可以保持其他信息要素的有效表现,包括创作时的原始版本、更丰富的语义格式和其他元素。本部分并不涉及可能会被嵌入在内的其他格式的长期适用性。

这些目标通过标识出在符合 PDF/A 标准的文件中可以使用的 PDF 组件集合以及对其使用形式的限制来实现。

PDF/A 本身并不一定确保内容的视觉外观能够准确反映用于创建合格文件的原始材料;例如,用于创建合格文件的流程可能会替换字体、重新调整文本流、采样图像或使用有损耗的压缩。需要确保合格文件准确表现原始材料的组织可能需要在本部分要求之外对生成合格文件的流程施加额外的要求,例如附录 C 中最佳实践所反映的情况。此外,为了能够对文件视觉外观是否正确进行检查确认,这些组织还应实施相关的制度和实用性措施。

PDF/A 并不直接涉及用于视觉表现的基础内容及 PDF/A 文件本身是否具有权威性的主题。这些权威性对于法律、监管和治理目的而言,普遍被认为是很重要的,但不在本部分的范围之内。

本部分是有利于组织机构长期保存电子档案的管理和技术环境的一个组成部分。组织机构能否成功实施本部分的内容,取决于其能否有效落实:

- 组织的档案环境的留存要求,GB/T 26162.1 中规定的记录管理制度和程序;
- 为确保电子文档及其特性长期留存而必需的额外要求和条件,包括但不限于 ISO 14721、ISO/TR 15801 和 GB/Z 23283—2009 所定义的要求和条件;
- 核实符合相关要求和条件所必要的质量保证流程,如核实已转换的来源数据的质量和完整性的检查制度。

本部分旨在促进开发与读取、呈现、生成和验证合格文件相关的各种应用。基于应用供应商理解的差异,不同的应用将会采纳本部分提出的不同功能要求,以准备、解读和处理合格文件。然而,这里应注意的是,合格的应用功能要能够适当地读取和处理合规范围内所有符合性级别的文件。

本部分扩展了第 2 部分的功能。和第 2 部分一样,基于 PDF 版本 1.7(定义见 ISO 32000-1)。

本部分(结合其规范性引用)为解释符合 PDF/A-3 标准的文件提供了充分的信息。

本部分的应用说明请参考 NPES 和 AIIM(均为被认可的标准制定机构)发布的最新信息。

文献管理

长期保存的电子文档文件格式

第3部分：支持嵌入式文件的 ISO 32000-1 的使用(PDF/A-3)

1 范围

GB/T 23286 的本部分规定了基于 PDF1.7 进行电子文档的长期保存,用于保持页面文档的静态视觉呈现,并允许包含任何类型的其他内容作为嵌入式文件或附件。

本部分不适用于:

- 将纸质或电子文档转换为 PDF/A 格式的具体流程;
- 具体技术设计、用户界面、实现方法或呈现的操作细节;
- 储存这些文档的具体物理方法,如媒介和储存条件;
- 需要的计算机硬件及/或操作系统。

注:本部分中的中英文词汇对照参见附录 NA。

2 规范性引用文件

下列文件对于本文件的应用是必不可少的。凡是注日期的引用文件,仅注日期的版本适用于本文件。凡是不注日期的引用文件,其最新版本(包括所有的修改单)适用于本文件。

ISO/IEC 646 信息技术 信息交换用七位编码字符集(Information technology—ISO 7-bit coded character set for information interchange)

ISO/IEC 10646 信息技术 通用多八位编码字符集(UCS)[Information technology—Universal Coded Character Set (UCS)]

ISO 15076-1 图像技术色彩管理 架构、配置文件格式与数据结构 第1部分:基于 ICC.1:2010 (Image technology colour management—Architecture, profile format and data structure—Part 1: Based on ICC.1:2010)

ISO 15930-7:2010 印刷技术 使用 PDF 进行印前数据交换 第7部分:使用 PDF1.6 进行打印数据的完整数据交换(PDF/X-4)与加入外部配置文件引擎的部分数据交换(PDF/X-4p) [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 文献管理 长期保存的电子文档文件格式 第1部分:PDF1.4 的使用(PDF/A-1) [Document management—Electronic document file format for long-term preservation—Part 1: Use of PDF 1.4 (PDF/A-1)]

ISO 19005-2 文献管理 长期保存的电子文档文件格式 第2部分:ISO 32000-1 的使用(PDF/A-2) [Document management—Electronic document file format for long-term preservation—Part 2: Use of ISO 32000-1 (PDF/A-2)]

ISO 32000-1 文献管理 可移植文档格式 第1部分:PDF1.7 (Document management—