



PDF Days Europe 2022 | Berlin

Navigating the PDF ecosystem

Peter Wyatt, PDF Association CTO

PDF Association technical resources

- <https://www.pdfa.org/resources/>

Technical resources

The following technical resource pages are actively maintained by the PDF Association to provide the PDF industry with the latest technical information.

Errata to ISO PDF related standards

This resource lists all resolved errata in PDF related ISO standards. It is presented as localized marked-up changes ([additions](#) and [deletions](#)) to the published wording in the relevant ISO standard organized by top level clause. Issue numbers (implemented here as popup tooltips) refer back to [closed GitHub pdf-issues with a "proposed solution" label](#). Due to ISO copyright, only minimal surrounding text from the relevant ISO standard is provided that is sufficient to locate where the resolution is being applied.

Status of PDF standardization at ISO

This resource summarizes the active and ongoing work items of all ISO committees dedicated to standardizing PDF. [ISO Technical Committee \(TC\) 171](#) Working Groups (WGs) are responsible for many PDF-related ISO publications. [ISO TC 130](#) WG 2 and related Task Forces (TFs) are responsible for PDF/X, PDF/VT, and other PDF-related graphic arts and commercial print applications. ISO TC 130 JWG 7 is the joint working group responsible for ICC and color management.

PDF specification index

This resource provides download links to legacy Adobe PDF references and errata, as well as the full ISO 32000 family of standards. It serves as an archive of the evolution of the core PDF file format.

All PDF related ISO Standards

The [index of all published ISO standards related to PDF](#) provides a mapping between ISO document numbers and the PDF nomenclature commonly used in industry (such as PDF/A, PDF/X, XMP, etc).

PDF 2.0 normative references

This resource provides download links for all Normative References of PDF 2.0 (ISO 32000-2:2020). Normative references are other additional standards and specifications where some or all of their content also constitutes technical requirements of PDF 2.0. They form essential reading for all developers of PDF 2.0 technology.

Guide to PDF Association Technical Working Groups

The PDF Association maintains various [interest-specific communities](#) facilitating members' discussions on technical issues related to PDF. If your interest is specific to Tagged PDF this [article](#) covers the focus on this area.

<https://www.pdfa.org/resources/>

Technical publications

The PDF Association provides many technical publications to assist developers and end-users to better utilize standardized PDF technologies.

Specifications

PDF Association specifications are industry-ratified specifications that extend PDF.

Title	Publication Date
PDF/Raster 1.0	July 20, 2017
PDF Declarations	September 5, 2019
Deriving HTML from PDF Version 1.0	June 11, 2019
EA-PDF	T.B.A.

Technical notes

Technical Notes provide additional clarity on the interpretation of ambiguous, unclear or purposely flexible statements in ISO PDF standards.

Title	Publication Date
Technical Note 0001: PDF/A-1 and Namespaces	March 14, 2008
Technical Note 0002: Color in PDF/A-1	March 14, 2008
Technical Note 0003: Metadata in PDF/A-1	March 18, 2008
Technical Note 0006: Digital Signatures in PDF/A-1	March 14, 2008
Technical Note 0008: Predefined XMP Properties in PDF/A-1	March 20, 2008
Technical Note 0009: XMP Extension Schemas in PDF/A-1	March 20, 2008
Technical Note 0010: Clarifications of ISO 19005, parts 1-3 for developers of PDF/A creators and validators	July 15, 2017
PDF 1.7 – ISO 32000-1 Summary of Changes	August 1, 2008

Other notes

[Adobe's patent release](#) regarding ISO 32000-1:2008 (2011 [statement](#) by Adobe's Matthew Hardy).

Application notes & best practice guides

Application Notes provide additional technical information around a specific feature of PDF that enables greater interoperability. Best practice guides offer concrete advice targeting specific outcomes.

Title	Publication Date
PDF 2.0 Application Note 001: Black Point Compensation	September 17, 2018
PDF 2.0 Application Note 002: Associated Files	October 26, 2018
PDF 2.0 Application Note 003: Use of object metadata streams	March 16, 2021
Tagged PDF Best Practice Guide: Syntax	June 10, 2019
PDF/VT Application Notes	January 3, 2016
Best Practice in Creating Print Files for Variable Data Printing - Designer Edition	June 2, 2022
Best Practice in Creating Print Files for Variable Data Printing - Developer Edition	June 2, 2022

PDF Association technical resources

- <https://www.pdfa.org/resources/>
- <https://www.pdfa.org/resource/pdf-specification-index/>
 - References to all old Adobe PDF references, errata, etc.

PDF Specification Index

This page provides external links to legacy Adobe PDF references and errata, as well as the ISO 32000 family of standards. The PDF Association is not authorized to host these files on behalf of the PDF industry.

This page was updated on 25 May 2022 as adobe.com has published a new resource for some legacy PDF reference documents at <https://opensource.adobe.com/dc-acrobat-sdk-docs/acrobatsdk/#pdf-reference>.

PDF Version	Year	Reference Documents
Adobe PDF 1.0	1993	<ul style="list-style-type: none">• Portable Document Format Reference Manual Version 1.0 (ISBN 0-201-62628-4)
Adobe PDF 1.1	1996	<ul style="list-style-type: none">• Portable Document Format Reference Manual Version 1.1 (ISBN 0-201-62628-4)
Adobe PDF 1.2	1996	<ul style="list-style-type: none">• Portable Document Format Reference Manual Version 1.2
Adobe PDF 1.3	2000	<ul style="list-style-type: none">• PDF Reference second edition, Adobe Portable Document Format Version 1.3• PDF 1.3 Errata (WayBack Machine)
Adobe PDF 1.4	2001	<ul style="list-style-type: none">• PDF Reference third edition, Adobe Portable Document Format Version 1.4• PDF 1.4 Errata (WayBack Machine)
Adobe PDF 1.5	2003	<ul style="list-style-type: none">• PDF Reference fourth edition, Adobe Portable Document Format Version 1.5• PDF 1.5 Errata (WayBack Machine)
Adobe PDF 1.6	2004	<ul style="list-style-type: none">• PDF Reference fifth edition, Adobe Portable Document Format Version 1.6• PDF 1.6 Errata (WayBack Machine)• Blend modes addendum (WayBack Machine)
Adobe PDF 1.7	2006	<ul style="list-style-type: none">• PDF Reference sixth edition, Adobe Portable Document Format Version 1.7

<https://www.pdfa.org/resource/pdf-specification-index/>

PDF Association technical resources

- <https://www.pdfa.org/resources/>
- <https://www.pdfa.org/resource/pdf-specification-index/>
 - References to all old Adobe PDF references, errata, etc.
- <https://www.pdfa.org/index-of-pdf-related-iso-publications/>
 - ISO publications related to PDF
 - Most ISO publications also have their own dedicated resource page

PDF = ISO 32000

The core PDF standard. A list of earlier Adobe PDF specifications and their errata can be found [here](#). A list of industry-based resolutions to [PDF 2.0 errata can be found here](#).

PDF version	Document
PDF 1.7	ISO 32000-1:2008 Document Management — Portable Document Format — Part 1: PDF 1.7
PDF 2.0	ISO 32000-2:2017 Document Management — Portable Document Format — Part 2: PDF 2.0 (withdrawn)
PDF 2.0	ISO 32000-2:2020 Document Management — Portable Document Format — Part 2: PDF 2.0

PDF/A = ISO 19005

PDF for long-term preservation and archival purposes, including [industry-based errata corrections for PDF/A-4](#).

Standard	Document	PDF version
PDF/A-1	ISO 19005-1:2005 Document management — Electronic document file format for long-term preservation — Part 1: Use of PDF 1.4 (PDF/A-1)	PDF 1.4
	ISO 19005-1:2005/COR 1:2007 Document management — Electronic document file format for long-term preservation — Part 1: Use of PDF 1.4 (PDF/A-1) — Technical Corrigendum 1	
	ISO 19005-1:2005/COR 2:2011 Document management — Electronic document file format for long-term preservation — Part 1: Use of PDF 1.4 (PDF/A-1) — Technical Corrigendum 2	
PDF/A-2	ISO 19005-2:2011 Document management — Electronic document file format for long-term preservation — Part 2: Use of ISO 32000-1 (PDF/A-2)	PDF 1.7
PDF/A-3	ISO 19005-3:2012 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)	PDF 1.7
PDF/A-4	ISO 19005-4:2020 Document management — Electronic document file format for long-term preservation — Part 4: Use of ISO 32000-2 (PDF/A-4)	PDF 2.0

PDF/E = ISO 24517

PDF for engineering. In PDF 2.0, PDF/A-4e has been specified to support engineering-centric 3D PDF workflows.

Standard	Document	PDF version
	ISO 24517-1:2008 Document management — Engineering document format using PDF — Part 1: Use of PDF 1.6	

PDF Association technical resources

- <https://www.pdfa.org/resources/>
- <https://www.pdfa.org/resource/pdf-specification-index/>
 - References to all old Adobe PDF references, errata, etc.
- <https://www.pdfa.org/index-of-pdf-related-iso-publications/>
 - ISO publications related to PDF
 - Most ISO publications also have their own dedicated resource page
- <https://www.pdfa.org/iso-status/>
 - ISO work-in-progress, updated after each ISO meeting

Contents

ISO TC 171 “Document management applications”

SC 2 “Document file formats, EDMS systems and authenticity of information”

- WG 5: “Joint TC 171/SC 2 - TC 42 - TC 46/SC 11 - TC 130 WG: Document management applications - Application issues - PDF/A”
- WG 7: “PDF/Engineering”
- WG 8 “PDF specification”
- WG 9 “PDF universal accessibility” (PDF/UA)
- WG 12 “Metadata”

ISO TC 130 “Graphic technology”

- WG 2 “Prepress data exchange”
 - TF 2: “Postpress” (PDF/X)
 - TF 3: “Variable data exchange” (PDF/VT)
 - TF 4: “XMP”
 - TF 5: “PDF Common Metadata”
- JWG 7: “Joint TC 130 - ICC WG: colour management”


TC 171 “Document management applications”

SC 2 “Document file formats, EDMS systems and authenticity of information”

Fall 2022: November 2022 (Virtual and In-Person - TBD)

WG 5: “Joint TC 171/SC 2 - TC 42 - TC 46/SC 11 - TC 130 WG: Document management applications - Application issues - PDF/A”

WG 5 is responsible for the ISO 19005 family of PDF/A standards for long-term preservation including archival applications.


ISO Number	Official title	Common name	Current status	Next key date	
ISO 19005-4	Document management — Electronic document file format for long-term	PDF/A-4	International Standard Published	Systematic Review 2025-12	PDF/A TWG

SC 2 “Document file formats, EDMS systems and authenticity of information”

Fall 2022: November 2022 (Virtual and In-Person - TBD)

WG 5: “Joint TC 171/SC 2 - TC 42 - TC 46/SC 11 - TC 130 WG: Document management applications - Application issues - PDF/A”

WG 5 is responsible for the ISO 19005 family of PDF/A standards for long-term preservation including archival applications.

ISO Number	Official title	Common name	Current status	Next key date	
ISO 19005-4	Document management — Electronic document file format for long-term preservation — Part 4: Use of ISO 32000-2 (PDF/A-4)	PDF/A-4	International Standard Published	Systematic Review 2025-12	PDF/A TWG
ISO 19005-1	Document management - Electronic document file format for long-term preservation - Part 1: Use of PDF 1.4 (PDF/A-1)	PDF/A-1	Confirmed for additional 5 years	Systematic Review 2025-12	PDF/A TWG
ISO 19005-2	Document management - Electronic document file format for long-term preservation - Part 2: Use of ISO 32000-1 (PDF/A-2)	PDF/A-2	International Standard Published	Systematic Review 2026-10	PDF/A TWG
ISO 19005-3	Document management - Electronic document file format for long-term preservation - Part 3: Use of ISO 32000-2 with support for embedded files (PDF/A-3)	PDF/A-3	International Standard Published	Systematic Review 2023-03	PDF/A TWG
ISO 18565 ¹	Document management - AFP/Archive	AFP/A	Confirmed for additional 5 years	Systematic Review 2025-12	PDF/A TWG
ISO 22550 ¹	Document management - AFP interchange for PDF	AFP	International Standard Published	Systematic Review 2026-12	PDF/A TWG
New Project	Document management - 2D Image file formats	2D	ISO Technical Specification Draft		PDF/A TWG
New Project	Document management - Text document long-term preservation		TBD		PDF/A TWG

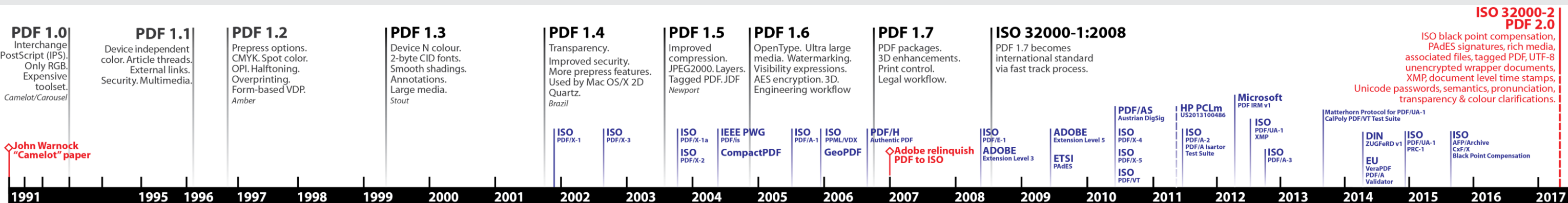
Before PDF...

- Adobe® PostScript®
 - Device independent page description language (vector graphics, fonts)
 - Dynamically typed, stack-based, concatenative programming language
 - Processed by an Interpreter (e.g. embedded in laser printer)
 - Always processed sequentially from start of file
- ‘Clone’ PostScript interpreters
 - Mimicked genuine Adobe PostScript implementation
 - Highly complex test cases and large test suites (e.g. Genoa) to achieve pixel accuracy



PDF background

- Adobe's John Warnock described "*The Camelot Project*" in early 1990s
 - Overcome PostScript limitations & weaknesses
- PDF is a random access binary object-oriented declarative page description language → "*digital paper*"
 - Defines graphical operations in a typeset color accurate transparent imaging model while relying on many nested formats for fonts, images, etc.



PDF Specification	Date	Pages	Random facts
Adobe PDF 1.0	June 1993	230	43 tables, 42 figures
Adobe PDF 1.1	January 1996	302	20 references
Adobe PDF 1.2	November 1996	394	137 tables, 86 examples
Adobe PDF 1.3	July 2000	696	223 tables, 73 figures
Adobe PDF 1.4	December 2001	978	277 tables, 20 color plates
Adobe PDF 1.5	August 2003	1172	333 tables, 70 figures
Adobe PDF 1.6	November 2004	1236	370 tables, 80 figures
Adobe PDF 1.7	October 2006	1310	389 tables, 98 figures
ISO 32000-1:2008 (PDF 1.7)	July 2008	756 (A4)	78 normative references
ISO 32000-2:2017 (PDF 2.0)	July 2017	984 (A4)	5868 “shall”, 410 “should” statements
ISO 32000-2:2020 (PDF 2.0)	December 2020	1000 (A4)	5891 “shall”, 414 “should” statements, 79 refs.

From Adobe to ISO

- Adobe proprietary (freely available) specification from PDF 1.0 – PDF 1.7
 - Less formal language
 - EBNF syntax up to PDF 1.2
 - Appendix H: Compatibility and Implementation Notes
- Adobe relinquished PDF 1.7 to ISO on 9 Jan 2007 for “fast track”(18 month) standardization process
 - ISO 32000-1:2008 (PDF 1.7) / equivalent available from Adobe (*free*)
- <https://www.pdfa.org/resource/pdf-specification-index/>
 - References to all old Adobe PDF references, errata, etc.

Changes in ISO 32000 over Adobe specs

- Use of formal standard-style language: shall, shall not, ...
 - Many ambiguous terms removed or resolved: should, may, might → can, ...
- International English (spelling, grammar)
- Updated and formalised all references
 - Latest RFCs, latest versions of other standards (CSS, XHTML), ...
- More cross-referencing within text
 - To other sections in PDF and to normative references
- Explicit description of some error and degenerate cases
- Requirement for consistency of redundant information
- Bringing in features from PDF/A, PDF/X and PDF/VT
- Use of “PDF reader”, “PDF writer”, “PDF processor” terminology
- Some Adobe implementation limitations now normative
 - e.g. { } depth on Type 4 PostScript Functions

PDF – Players

Organisations

(private, NFP, government, NGO)



Associations

(industry, profiles)



International Standards



PDF standards & specifications

ISO

Current industry

Specialized

Graphic Arts &
Commercial Print

GWG

Print product
metadata

Processing Steps

PDF/VCR

PDF/VT

PDF/X

Archival & Preservation

Engineering

GeoPDF

VDA

PDF/E

PDF/AS

PDF/mail

PDF/A

e-Finance

...

Chile DTE PDF

Brazil RPS PDF

Order-X

Fractur-X

ZUGFeRD

Accessibility
& Reuse

Deriving HTML
from PDF

Well-Tagged PDF

PDF/UA

Other industries...

Healthcare

ICH (pharma)

PDF/H

...

Imaging/IoT

PDF/raster

PDF/is

PDF/R

ISO 32000 & Adobe (legacy) Base Versions

ECMAScript for PDF

PDF Declarations

ISO/TS Extensions

RFC

XFDF

...

PRC

ISO Normative References

- ISO Normative References

ISO boilerplate: *“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 or corrigenda) applies.”*

- **Dated** normative reference → *exactly and only as stated*
- Undated normative reference → latest version, possibly with errata
Discussion: *Does this mean PDF is a moving target or just reflecting reality?*
- <https://www.pdfa.org/iso-32000-normative-references/>
 - Additionally includes URLs and informative notes (*not allowed by ISO Directives*)
- Bibliography is always informative (just “FYI”)

Importance of normative references

- Normative references have their own normative references
 - Creates a “tree” of normative dependencies
 - These dependencies can **conflict/disagree!**
- <https://github.com/pdf-association/PDF2NormRefs>
- <https://safedocs.pdfa.org/> - Normative References flythrough

Hands-on / Discussion

- Unicode for passwords vs Unicode in content
<https://github.com/pdf-association/safedocs/tree/main/Unicode%20passwords>

Current ISO activities



- <https://www.pdfa.org/iso-status/>
 - ISO TC 130 WG 2: October 2022
 - PDF/X, PDF/VT, Print Product Metadata, ICC
 - ISO TC 171 SC 2: November 2022
 - PDF, PDF/A, PDF/UA, ECMAscript, various extensions & clarifications, PRC
 - Previously: PDF/E, PDF/R, AFP
- ISO document drafts available to PDFa members
 - ***Great source of information!***
 - Can review & submit comments – *or just stay informed*
 - <https://www.pdfa.org/iso-n-documents/>

Current ISO work items

- **ISO 32000-2:2020** (PDF 2.0) Amendment 1
 - Errata via XFDF
- **PDF/A** [TC 171] and **PDF/X** [TC 130]
 - Errata and issues for discussion – *no decision is made*
- **PDF/UA-2**
 - Separation of file format and processor requirements
- **New work:** 2D image codecs under early discussion
 - New image compression codecs, HDR, ...
- **ECMAScript for PDF 2.0**
 - Errata and issues for discussion – *no decision is made*



PDF 2.0 (ISO 32000-2)

- PDF 2.0 took 9 years to develop under full ISO open collaborative process
- ISO 32000-1:2008 → ISO 32000-2:2020 (PDF 2.0)
 - Corrects and clarifies many aspects of (legacy) PDF
 - Adds new features, some features marked as “deprecated” (§3.15)
 - Nothing removed (“obsoleted”)
- **If it was previously ambiguous, refer to ISO 32000-2:2020 (PDF 2.0)!**
- Errata process via <https://github.com/pdf-association/pdf-issues/>
 - Industry-agreed resolutions at <https://pdf-issues.pdfa.org/>

PDF subsets

- PDF “base versions” define all syntax, rules & requirements for all uses
 - “*Everything is optional*” → highly permissive and thus flexible
 - Adobe PDF 1.x → ISO PDF 2.0 (ISO 32000-2)
- **Subsets** support specific industries / use cases
 - Derive from a *specific* PDF base version
 - Mandate, prohibit or extend
 - But cannot contradict PDF base version requirements
 - Self-declared conformance (via XMP)
 - Stricter rules allow consistent detailed validation (“pre-flight”)
 - Common wording across most subsets now with PDF 2.0 subsets

ISO TS extensions – in progress

ISO Technical Specification (TS)	New feature?	PDF Association Community / ISO TC 171 SC 2 WG
ISO/TS 32001 Document management - Portable Document Format - Encryption and Hash algorithm support in ISO 32000-2 (PDF 2.0)	Yes	Digital Signature TWG WG 8
ISO/TS 32002 Document management - Portable Document Format - Extensions to Digital Signatures in ISO 32000-2 (PDF 2.0)	Yes	Digital Signature TWG WG 8
ISO/TS 32003 Document management - Portable Document Format - Adding support of AES-GCM in PDF 2.0	Yes	Digital Signature TWG WG 8
ISO/TS 32004 Document management - Portable Document Format - Integrity Protection in encrypted document in PDF 2.0	Yes	Digital Signature TWG WG 8
ISO/TS 32005 Document management - Portable Document Format - PDF 1.7 and 2.0 namespace inclusion in ISO 32000-2	No	PDF/UA TWG WG 9
ISO/TS 24654 Document management – Portable Document Format – Non-Rectangular Links	Yes	PDF TWG WG 7
ISO/TS 24064 Document management – Portable Document Format – 3D data streams conforming to the ISO 10303:242 (STEP AP242) specification	Yes	3D PDF TWG WG 7
ISO/TS 6912 Document management — Portable Document Format — Clarification for initial graphics state in ISO 32000-2 (PDF 2.0)	No	PDF TWG WG 8
ISO/PDTS 32007 Document management – Portable Document Format – RichMedia annotations conforming to glTF assets	Yes	3D PDF TWG WG 7

PDF standards and ISO terminology

- Set by ISO/IEC Directives
 - Part 1: Procedures for meetings
 - Part 2: Writing rules
- Normative vs. informative
 - Informative: some annexes, all notes, all examples
- “Shall” vs. “Should” (vs. statement of fact)
 - No order or hierarchy of rules unless explicitly stated
- File format vs. Processor requirements
 - Reader vs Writer. Interactive vs non-interactive.
- “Key” vs. “value” vs. “entry”

The reality

- The core PDF specification has evolved over time
 - Usually for the better
 - Corrections, clarifications, addition of missing details, internal cross-referencing, vendor neutral, level playing field, consistency of terms, ..., ...

There is no “track changes” version.

Use of more formal language (“ISO-ese”).

Some sections have been entirely rewritten, others are unchanged.

Section ordering is largely unchanged.

Small text change → possibly **big** technical change!

No hierarchy of rules

<< /Foo null /Foo /Bar >>

- Clause 7.3.7 Dictionary objects:
“Multiple entries in the same dictionary shall not have the same key.”
- Clause 7.3.9 Null object:
“Specifying the null object as the value of a dictionary entry (7.3.7, “Dictionary objects”) shall be equivalent to omitting the entry entirely.”

File or processor requirement?

- “.... shall be ignored if ...”
- “... shall [not] be present if ...”
- “... shall [not] be used ...”
([PDF/A TWG Issue #17](#))
- “... shall [not] be absent ...”
- “... is meaningful only if ...”
([Errata #6](#))

Is there any difference if a requirement refers to a key, a value, or an entry?

What if the values' object type of a key is not correct?

PDF key trends

- Majority of users have no idea what product they use to view PDF
- More PDF software = **more variation in appearance, behaviour & support**
More PDF software = **more ‘bad PDFs’**
 - Threat to “accurate & reliable digital paper”?
- Uncommon to publish any compliance statements
 - Microsoft Edge previously published a **compliance statement**
 - Others do not:

Handling of PDF versions

The current implementation is a mix of PDF 1.4 and some adhoc additions without a clear distinction what is and is not supported. We could add some support for explicitly handling versions in PDFBox e.g. by marking certain methods and properties to the PDF version support level. This could in addition be a good basis for PDF/A and other compliance checks.

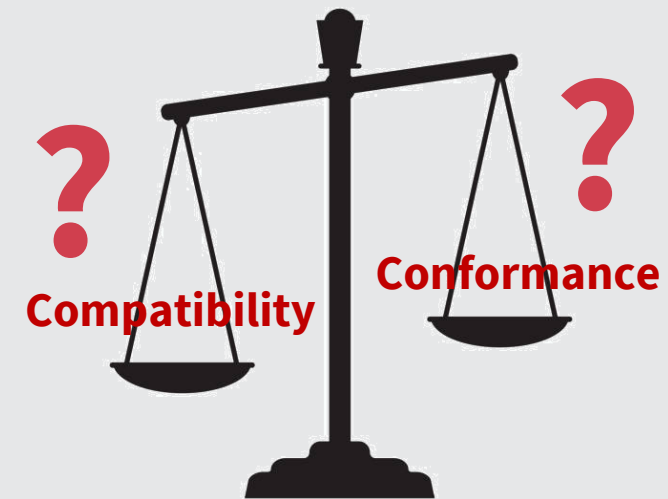


2.1.8 [ISO32000-1] Section 8.7.4.5.2, Type 1 (Function-Based) Shadings
V0008:
The specification states:
<div>In Type 1 (function-based) shadings, the colour at every point in the domain is defined by a specified mathematical function.</div>
Microsoft Edge:
Function-based shadings are not supported (their presence is ignored).
2.1.9 [ISO32000-1] Section 8.7.4.5.5, Type 4 Shadings (Free-Form Gouraud-Shaded Triangle Meshes)
V0009:
The specification states:
<div>Type 4 shadings (free-form Gouraud-shaded triangle meshes) are commonly used to represent complex coloured and shaded three-dimensional shapes.</div>
Microsoft Edge:
Free-form Gouraud-shaded triangle mesh shadings are not supported (their presence is ignored).
2.1.10 [ISO32000-1] Section 8.7.4.5.6, Type 5 Shadings (Lattice-Form Gouraud-Shaded Triangle Meshes)

The reality

- Many PDF files do not comply with the PDF specification (a.k.a. errors)
 - Some are minor, others are more major. Range from fatal to not noticeable.
 - Some are recoverable, some are not (*depending on effort*)
 - PDF files persist forever → some “well known errors”

- Errors can occur anywhere: lexical, syntax, file, DOM, ...
- **Conformance:** compliance to the specification
- **Compatibility:** PDFs Just Work™ (*no matter how invalid?*)



The reality

- PDF software is permissive to some degree
 - Explicit: coded
 - Implicit: because a 3rd party library or operating system is used
 - Because the software never reads a certain PDF data structure and doesn't notice
- ***Realistically:*** what can users do about errors in their PDFs?
- *What can PDF developers do about errors in PDFs they find?*

- **Permissiveness (leniency):** allowing malformations, errors, etc.
- **Differential:** different outcomes from different software
- ***Just because “it works” in a viewer does not mean it is correct!***

The reality

- PDF files created using software written with older specifications:
 - ... may be invalid against a newer specification.
 - ... might produce different results against a newer specification.

Examples: resource inheritance, deprecated features, standard 14 fonts, font descriptors, ...

Facts of (PDF) Life

- PDF files can be malicious
 - Intentional vs. unintentional
- Malicious PDFs can be 100% valid
 - Benign PDFs can be invalid or malformed
- PDF is a complex
 - PDF software is highly complex, often includes many 3rd party software dependencies
- Attackers are increasingly sophisticated
 - Social engineering, phishing, supply chain, multi-platform, ...
- Most users of PDF are not technical
- **Protecting users = protecting the brand of PDF = helping our industry**



Evil hackers always wear hoodies

Security & PDF

- Commonly utilised PDF features for attacks
 - JavaScript, actions, URL, form submit, screen control, embedded files
- JavaScript (ECMAScript) has been biggest source of:
 - F.U.D., real-world and PoC exploits
 - Research by security hackers & academics
- Newer research highlights other aspects...

“VirusTotal receives over 12 million (non-executable) document submissions per year”

“A Broad View of the Ecosystem of Socially Engineered Exploit Documents”, Blond et al, 2017.