



Dietrich
von Seggern

Managing
Director
callas
software

Vice Chair PDF
Association



electronic
document
CONFERENCE

PDF's ISO-standardized subsets: a tour





PDF 2.0



Current: PDF 2.0 - ISO32000:2017

- Developed in almost 10 years, almost 1000 pages
- Completely developed at ISO
 - Adobe: PDF 1.0 - 1.7
 - ISO: PDF 32000-1 and 32000-2
(ISO 32000-1 almost identical to PDF 1.7)



Next: PDF 2.0 - ISO32000:2020

- ISO is working on a “dated revision”
- No new features, mostly bug fixing
- Just sent to DIS (5 months ballot)
- Probably published in 2020
- *Will be the basis for new parts for all PDF based subset standards*



Why PDF subset standards?

- PDF is powerful and complex
 - The 1000 pages reference several 1000 pages of related specifications (Fonts, Unicode, ICC profiles, XMP, ...)
- PDF is everywhere and essential for many workflows (archival, print, ...)
- The *purposes* of PDF vary and wherever PDF is essential it is important that specific requirements are met



General goals for PDF subset standards

- Dedicated standards for distinct processes
- Clear and simple rules
 - For internal PDF based processes
 - For interoperable processes between organizations
- Reliable technical foundation, relieving users of technical details
- Clarifying responsibilities between creator and processor



Overview

- PDF/X
- PDF/A and PDF/E
- PDF/UA
- PDF/VT and PDF/VCR
- PDF/R



Why not just one standard for “reliable PDF”?

- Dedicated standards for distinct processes
 - Requirements for reliable device independent print (PDF/X) are different from reliable screen rendering (PDF/A-2b)
 - Requirements for PDF/UA (accessibility) are more demanding than for rendering (PDF/A-2b)
- The next generation of PDF based subset standards will use the same text wherever possible
- Making it easier for creators to comply with more than one standard





PDF/X



PDF/X - fact sheet

- Oldest PDF based Standard (1999/2001)
- EXchange of pre-press files, pre-product for print
- Closely related to print standards, e.g. ISO 12647 for offset
- Widely used and accepted
- Covers all printing methods, therefore it also very broad, became basis for industry standards specializing on print methods:
 - Ghent PDF Workgroup
 - PDFX Ready



PDF/X - standard parts

PDF/X-1a (2001)

- CMYK
+ spot colors
- No transparency, no layers
- Widely used, straightforward

PDF/X-3 (2002)

- CMYK
+ spot colors
+ ICC (device independent)
- No transparency, no layers
- More flexible (repurposing)
- Higher requirements for processes and devices

PDF/X-4 (2010)

- CMYK
+ spot colors
+ ICC (device independent)
- Transparency and layers
- Devices that can deal with “live transparency”
- Basis for PDF/VT and PDF/VCR



Next: PDF/X-6 = PDF/X-4 + ...

- Annotations permitted if “appearance” is defined
- PDF 2.0
 - Page based Output Intents
 - Black Point Compensation
 - Halftone Origin
 - Spectral measurements for spot colors (CxF)
 - DPart metadata (for pages)
- Is going to DIS (in parallel with PDF 2.0 dated revision)



Alternatives?

- Not really
- “Open files” (InDesign, Quark XPress, Office,...)
- Exception: **Customer communications**
“AFP” - in many cases a container for PDF
(standardized as “AFP interchange for PDF” ISO 22550)





PDF/A and PDF/E

PDF/A - fact sheet

- PDF/A-1 was published in 2005
- Requirements for archival in industry and governments, driven by US organizations
- First used mainly in Europe to replace TIFF in archives
- Today for mostly digitally created documents (Digital Office)
- Reliable document exchange between organizations (e.g. ZUGFeRD invoices)
- Not just pre-product for print - as PDF/X (but compatible)



PDF/X - standard parts

PDF/A-1b

- Forms and annotations: Requirements apply
- Compression: no LZW
- Fonts have to be embedded
- Colors have to be device independent
- Requirements for metadata based on XMP

PDF/A-2b

- Plus
- JPEG2000
 - Transparency
 - Layers

PDF/A-2u

- Unicode (text search)

PDF/A-1a + PDF/A-2a

- Unicode (text search)
- Tagging structure



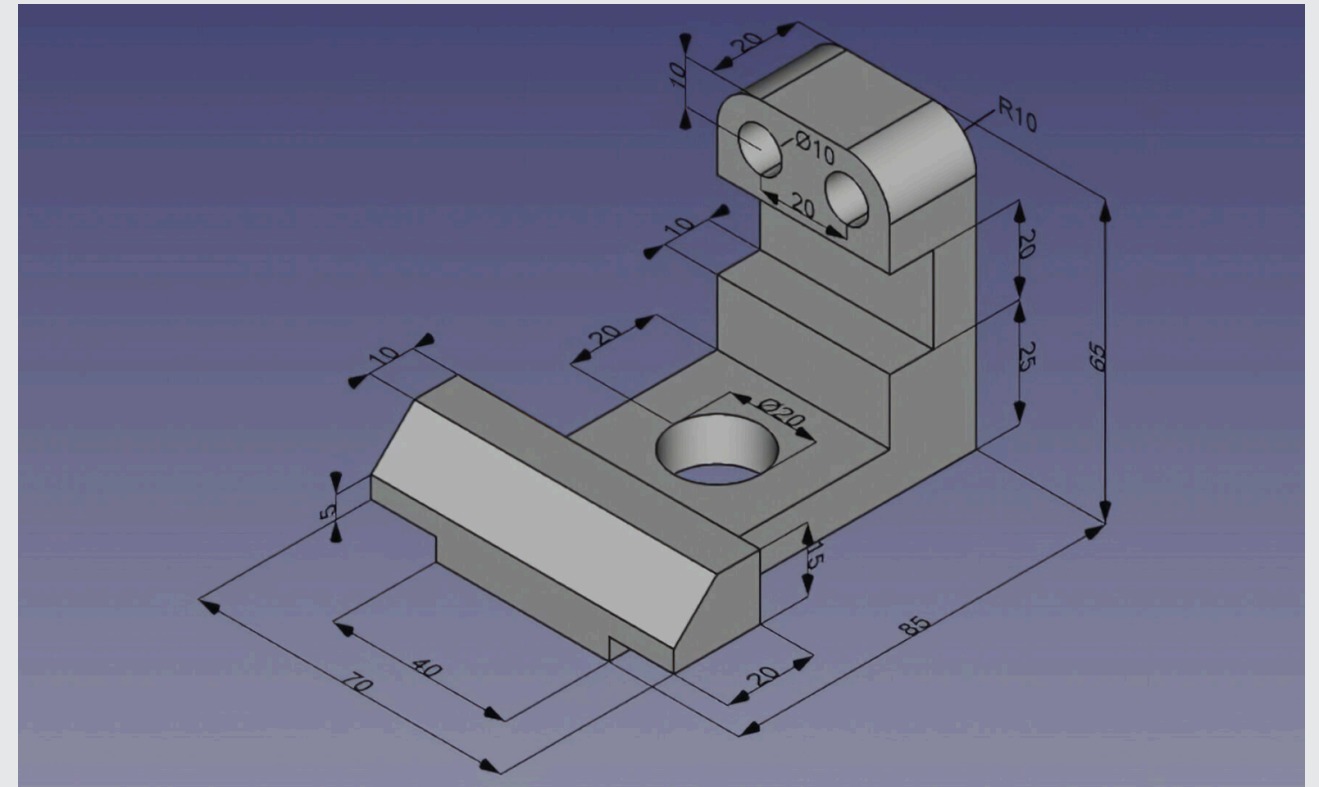
Next: PDF/A-4, PDF/A-4e and PDF/A-4f

- Relaxed requirements for metadata and font subsets
- Based on PDF 2.0
 - Current state-of-the-art signatures and encryption
 - Improved Tagging
 - PDF/A-4 as successor of PDF/A-2
 - PDF/A-4f as successor of PDF/A-3
 - PDF/A-4e as successor of PDF/E...
- Is going to DIS (in parallel with PDF 2.0 and ...)



PDF/E - fact sheet

- Published in 2008
 - Exchange format for 3D drawings
 - Almost no practical use
- Industry has an interest in a standardized archivable format
 - Car manufacturers standardize using PDF/A-3 for that purpose
- PDF/A-4e
 - 3D-Formats: U3D, PRC



Alternatives?

- **Not** really, **Paper** or **TIFF**?
- Problematic are video, audio, websites etc.
where PDF is not always ideal





PDF/VT & PDF/VCR

PDF/VT - fact sheet

- Published in 2010
 - PDF/VT-1: Complete files (thousands of pages)
 - PDF/VT-2: Referenced (variable) objects
- Designed as exchange format for variable and transactional print
- Based on PDF/X-4
- Alternative to AFP, (PCL, PPML etc.)
 - Enhanced graphical model
 - Widely available viewers
- Not very much in use...



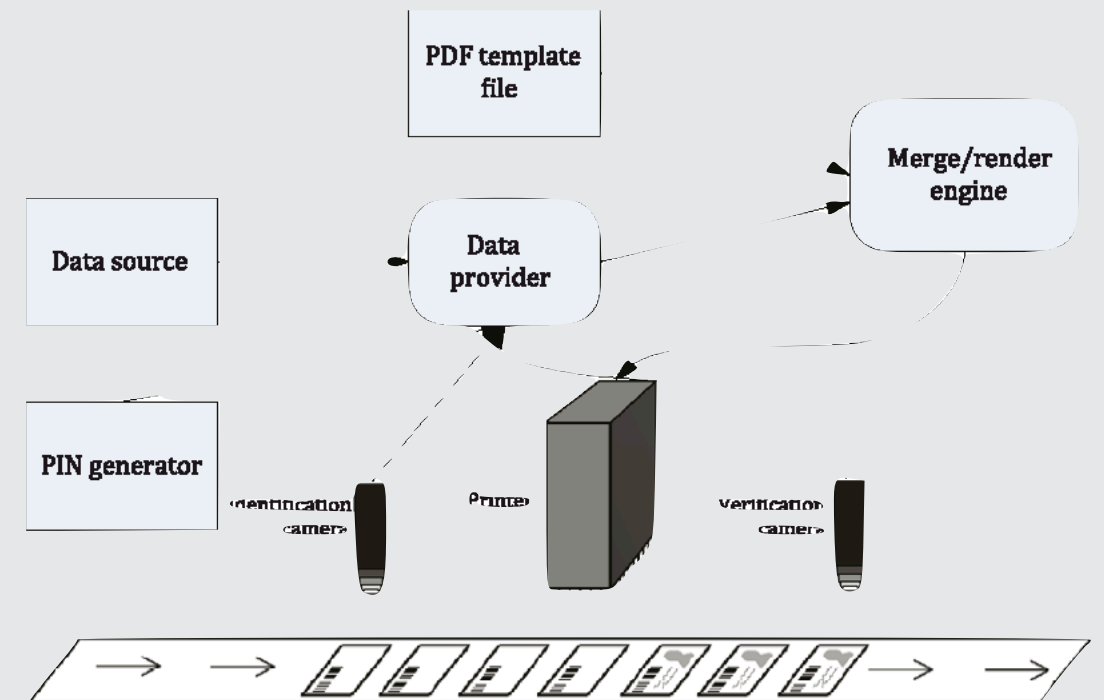
PDF/VT - problems and solutions

- Specifies page based metadata, but only syntax, no semantics
 - No interoperability for printer control (e.g. paper tray selection)
 - No interoperability for finishing (e.g. inline binding)
- Closing the gap: “Print product metadata for PDF files” (ISO 21812-1) currently in FDIS
- Current ISO project: PDF/VT-3
 - Based on “PDF/X-6” and “Print product metadata for PDF files”
 - Is going to DIS (in parallel with PDF 2.0)



PDF/VCR - fact sheet

- Published in September 2017
- Driven from existing implementations
- Similar scope as PDF/VT
- For “real time” printing, e.g. read data from a credit card and use it when printing the envelope



Alternatives?

- AFP - possibly as a container for PDF
(standardized as “AFP interchange for PDF” ISO 22550)





PDF/UA



PDF/UA - fact sheet

- Published 2014 (UA = universal accessibility)
- Accessibility “baked in” into HTML and eBooks formats
- In PDF structure optional via “tagging” structure
- Accessible documents are required by law in restricted industries (public sector, insurance companies etc.)
- Tagging is the basis for content repurposing and text reflow



PDF/UA - technicalities

- Technical basis for unicode is PDF/A-2u
- Syntactical and semantical requirements for tagging structure
 - Tags similar to HTML
 - Headlines, paragraphs, columns, tables ...
 - Alternate text for images and diagrams
- Automated tagging in PDF not permitted, because not reliably possible
- Creation and export from the authoring applications (editorial systems, office or layout applications)



Next and Alternatives?

- PDF/UA-2 needs more time than the other PDF based standards (currently in CD)
- Alternative: **HTML, EPUB**



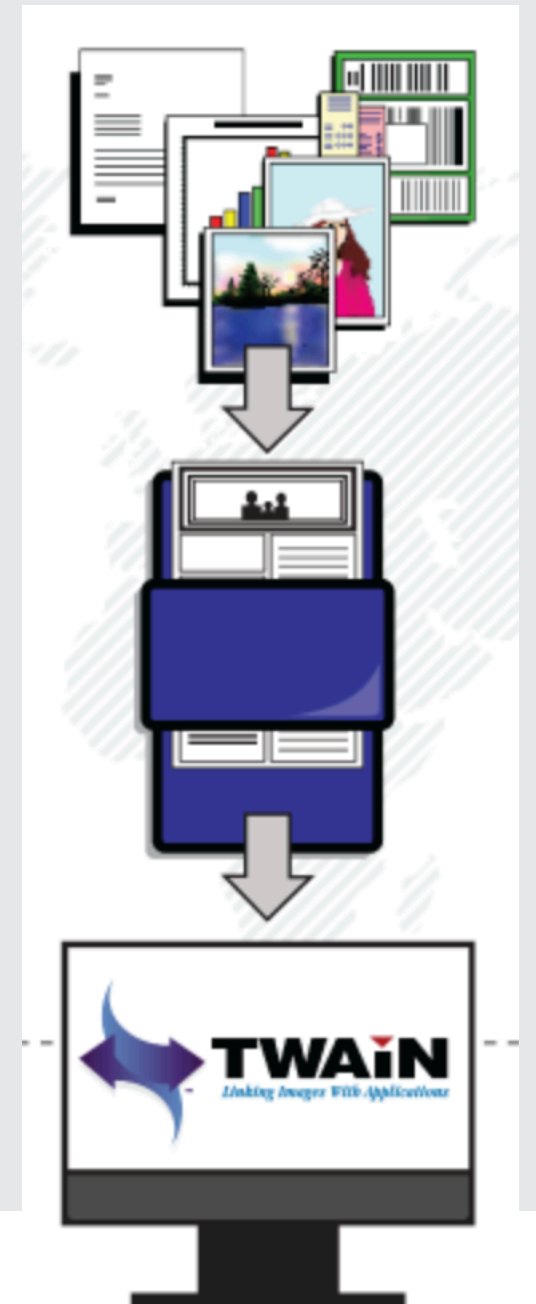


PDF/R



PDF/R

- Based on PDF/raster, a standard developed by TWAIN and PDF Association (*TWAIN represents the imaging industry and develops standards that link applications and image acquisition devices*)
- Replacing TIFF/JPEG in the interface between scanner and workflow
 - Multiple pages
 - Better compression
 - Better, widely available viewers
- Severely limited set PDF features, basically only one-, two or RGB colored pages
- Going to DIS (in parallel with PDF 2.0)



Alternatives?

- Not really, TIFF?



Overview



Standard	Published	Scope	Use
PDF/X	2001	pre-product for print	very high
PDF/A	2005	digital archival	high
(PDF/E	2008	engineering	none)
PDF/VT	2010	variable, transact. printing	low
PDF/UA	2014	accessibility, richer docs	some
PDF/VCR	2017	variable print in real time	specialized
PDF/R	2020?	scanner to workflow interface	probably high





Questions?

Comments are welcomed.



Thank you!

Dietrich von Seggern

Managing Director callas software

Vice Chair PDF Association

We appreciate your participation.



Thank you!

Any questions?

Get in touch: dietrich.seggern@pdfa.org
Web site: www.pdfa.org
Twitter: PDFAssociation

