

Developing PDF

Matthew Hardy
Sr Engineering Manager, Adobe



Matthew Hardy,
Sr Engineering Manager
Adobe



About Me

- Senior Engineering Manager for Adobe Document Cloud
- 12½ years at Adobe
- ISO representative for PDF standards since 2010
- Convener for WG8 (PDF)
- ISO project leader for ISO 19444 (XFDF)
- Standards Board member for 3D PDF Consortium
- PhD in document semantics with focus on PDF



Matthew Hardy,
Sr Engineering Manager
Adobe



Overview

- Road to PDF 2.0
- PDF 2.0 based specifications
- Other standards
- Building on PDF 2.0
- PDF 2.0+ active development
- Future of PDF



Matthew Hardy,
Sr Engineering Manager
Adobe



Road to PDF 2.0

- PDF 1.0 software released in 1993
 - RGB, Links (internal), Bookmarks, Fonts, etc.
- PDF 1.1 - 1.7 introduced through 2006
 - Updates ~2 years apart
 - Many new features
 - CMYK, ICC Colors, Commenting, Security, Digital Signatures, Forms, Tagged PDF, 3D and much more...
- ISO 32000-1:2008 based on PDF 1.7
- ISO 32000-2:2017 defined PDF 2.0
 - Namespaces for Tagged PDF, RichMedia (3D enhancements), GeoSpatial, Long Term Validation for DigSig, and much more...



Matthew Hardy,
Sr Engineering Manager
Adobe



Subsets: PDF/A and PDF/E

- PDF/A-4 being developed based on PDF 2.0
 - Page-level output intents
 - Associated file enhancements
 - Tagged PDF and Namespaces
 - Scripting
- PDF/A-4f conformance level
 - Unifying PDF/A-2 & A-3 embedded files with conformance levels
- PDF/A-4e conformance level
 - Bringing engineering archival (PDF/E concepts) to PDF/A
 - Effectively, PDF/E-2 merged into PDF/A-4e
 - 3D now permitted (U3D and PRC)



Matthew Hardy,
Sr Engineering Manager
Adobe



Subsets: PDF/UA

- PDF/UA-2 being developed based on PDF 2.0
 - PDF 2.0 Tagged PDF support
 - PDF 1.7, 2.0 and MathML structure types
 - Namespaces
 - Associated files for MathML
 - Focus on tagging what's present
- Extending PDF 2.0
 - Considering extended tagging
 - ARIA-attributes for deeper semantics
 - Feedback into PDF 2.0+



Matthew Hardy,
Sr Engineering Manager
Adobe



Subsets: PDF/X & PDF/VT

- PDF/X-6 being developed based on PDF 2.0
 - Per-page Output Intents
 - DPart Metadata
 - Black Point Compensation
 - Halftone Offset Specifications
 - Removes forms restrictions
- PDF/VT-3 also being developed based on PDF 2.0
 - Utilizes features of X-6 for variable and transactional printing



Matthew Hardy,
Sr Engineering Manager
Adobe



Subsets: PDF/raster

- PDF/raster is a PDF Association standard
 - Partnership with TWAIN
- Currently in process to become ISO standard
- Highly restrictive subset of 1.7/2.0
 - Easy to create and process on devices
- Targeted at the scanner industry (TWAIN)
- Leverages new encryption in PDF 2.0



Matthew Hardy,
Sr Engineering Manager
Adobe



Other Standards: XFDF

- ISO 19444-1:2016 defines the XML Forms Data Format
- Based on ISO 32000-2:2017
- XFDF transports annotation data
- Dated revision
 - Extending PDF 2.0 annotation support
 - RichMedia and 3D focus



Matthew Hardy,
Sr Engineering Manager
Adobe



Other Standards: PRC

- ISO 14739-1:2014 defines PRC
 - One of two 3D formats supported in PDF 2.0
- Dated revision
 - Correct errata found in spec
 - Improved descriptions of algorithms
 - Increased interoperability
 - Improve PDF 2.0 support



Matthew Hardy,
Sr Engineering Manager
Adobe



Building on PDF 2.0: ZUGFeRD

- Utilizes PDF/A-3 Associated Files
 - Now part of PDF 2.0 ISO 32000-2:2017
- Classic PDF
 - Represents 2D graphical representation of invoice
 - Viewable in any PDF viewer
 - XMP indicates ZUGFeRD
- XML Associated File
 - Attached at document level
 - Core Cross Industry Invoicing XML standard
 - Relationship *Alternative*
 - Naming convention



Matthew Hardy,
Sr Engineering Manager
Adobe



Building on PDF 2.0: MathML

- Utilizing Associated Files
 - PDF/A-3 & PDF 2.0
- Rich data extraction for MathML
- Re-visualization of Mathematics
 - Increase interactivity
 - Guides
 - Accessibility
 - Re-use



Matthew Hardy,
Sr Engineering Manager
Adobe



Building on PDF 2.0: Open Data

- PDF open data
 - W3C PDF and Open Data Community Group
- Rich open data within PDF
 - Sources
 - Data
 - Schema
 - Alternatives
- Participate in the rich open data community
 - Archive
 - Connectivity
 - Workflows



Matthew Hardy,
Sr Engineering Manager
Adobe



Building on PDF 2.0: Content Re-use

- Tagged PDF & Namespaces
 - Enable richer communication of semantics
 - Content re-use
 - Accessibility
 - Repurposing
- Re-use
 - Extract to source
 - Extract HTML (e.g. Callas PDFGoHTML)
 - Adapt to different screens and environments
 - Richer re-purposing



Matthew Hardy,
Sr Engineering Manager
Adobe



Building on PDF 2.0: PDF as a Package/Container

- PDF is a container
 - Visual content
 - Related materials
 - Data
 - Semantics
 - Metadata
- Package
 - Deliver a self-describing experience
 - Rich archive for all related material
 - Documents, email and more



Matthew Hardy,
Sr Engineering Manager
Adobe



PDF 2.0+ Active Development

- Dated Revision for Errata in ISO 32000-2:2017
 - Minor mistakes and typos
 - Plan to keep changes editorial
- PDF 2.1 and beyond
 - Encryption and signing (AES-256 GCM & PKCS#1 2.2)
 - Non-rectangular annotations (e.g. links)
 - Variable width Bezier curves
 - STEP 3D
 - ICCNBased colors
 - Tagged PDF gaps



Matthew Hardy,
Sr Engineering Manager
Adobe



PDF at 50

- 25 years
 - Lot of evolution of the format
 - Started very static
 - Now very interactive
- 25 more years...
 - Blockchain
 - Richer data
 - Greater semantics
 - Reimagine PDF as a dynamic format
 - Plenty of life in the format!



Matthew Hardy,
Sr Engineering Manager
Adobe



Thank you!

Any questions?



Matthew Hardy,
Sr Engineering Manager
Adobe

Get in touch:

Web site:

Twitter:

matthew.hardy@adobe.com

<http://www.adobe.com/>

[@mrbhardy](https://twitter.com/mrbhardy)

