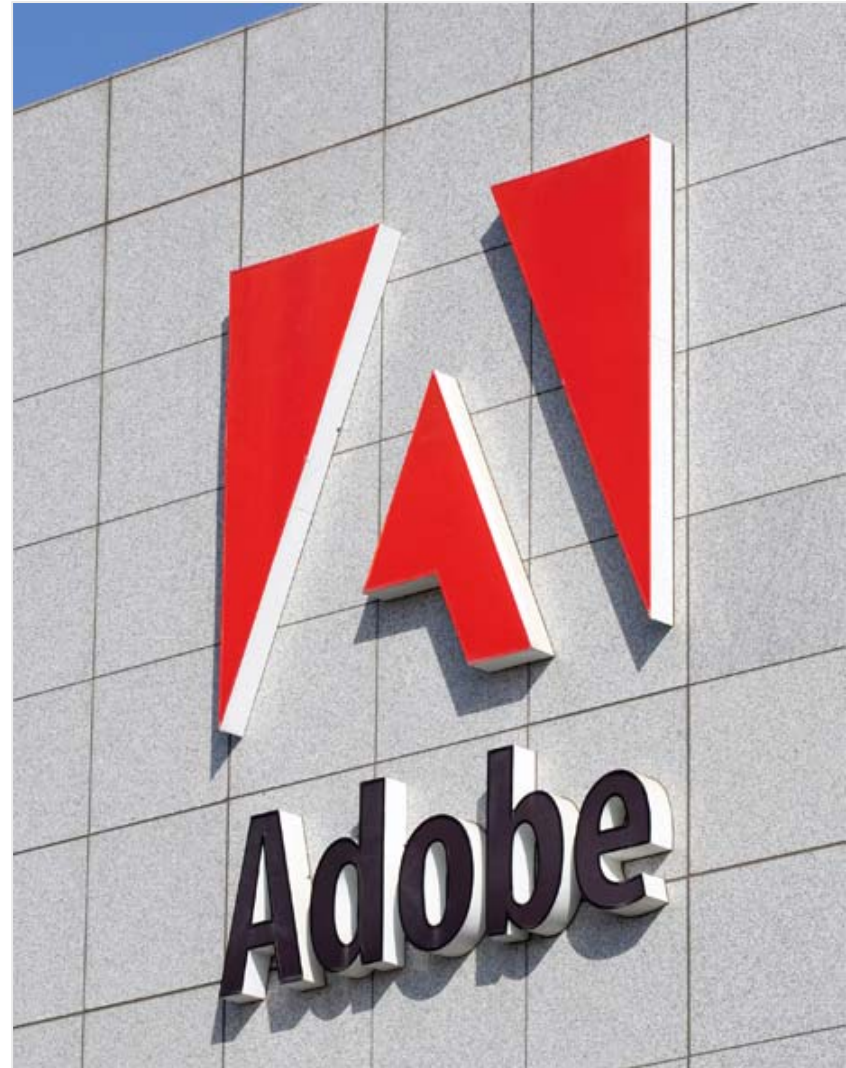


# Archiveren met PDF/A (een introductie)

Colin van Oosterhout

Business development manager Acrobat

Adobe Systems Benelux



# Vier aspecten van digitale documenten

- **Het begrip “digitaal document” als een combinatie van hardware, software en computerbestanden**

- Een computerbestand kan alleen gelezen worden via een specifieke combinatie van hardware en software, bijvoorbeeld Word voor de apple macintosh

- **Metadata ter veiligstelling van authenticiteit bij digitale documenten**

- Vele soorten, waaronder :  
beheergegevens en institutionele context

- **Digitale kenmerken**

- Inhoud, structuur, context, presentatievorm en gedrag

- **Authenticiteit bij digitale documenten**

- Het document is wat het claimt te zijn en is niet veranderd (integriteit en verificatie spelen een kernrol)



# De archiefregeling uit 2001

Regeling van de Staatssecretaris van Onderwijs, Cultuur en Wetenschappen, dr. F. van der Ploeg, van ..... 2001, nr. WJZ/2001/..... (8080), houdende nadere regels omtrent de geordende en toegankelijke staat van te bewaren archiefbescheiden. Gelet op artikel 12 van het Archiefbesluit 1995,

## Artikel 6

**Digitale archiefbescheiden dienen, uiterlijk op het tijdstip van overbrenging, als bedoeld in de artikelen 12 en 13 van de Archiefwet 1995, te worden opgeslagen volgens de**

**Volgende standaarden:**

- a. voor character sets: ASCII (ISO/IEC 8859-1) of Unicode (ISO/IEC 10646-1);**
- b. voor tekstbestanden: Portable document format (PDF) of SGML dan wel XML vergezeld van een stylesheet (XSL, CSS) dan wel TIFF of PDF met de metadata in een XML-wrapper;**
- c. voor CAD/CAM bestanden; Portable document format (PDF) en STEP (Standard for the exchange of product data) als metadata standaard (ISO 10303);**
- d. voor images/beelden (bitmapped): Portable document format (PDF) en indien gebruik gemaakt wordt van compressie: ITU T4 of ITU T6;**

# Genoemde formaten uit de archiefregeling

- ASCII
- SGML
- XML
- PDF
- TIFF



# Bijvoorbeeld over ASCII

- Omzetting van bit naar karaktertekens
- American Standard Code for Information Interchange
- Oorspronkelijk 7 bits (128 karaktertekens), later 8 bits dus 256 karaktertekens. Ook dit was te weinig (bijvoorbeeld voor de Griekse taal) en (mede) daarom is unicode ontwikkeld

# Bekende Archiveringsstrategieën

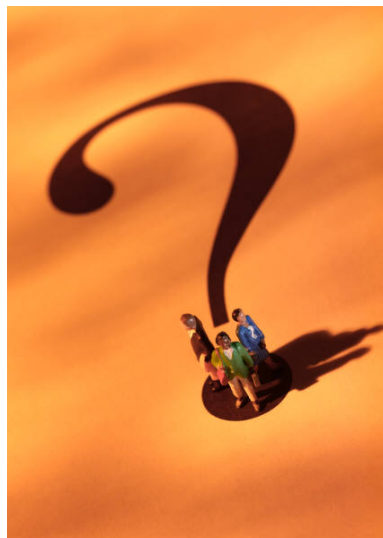
## ▪ Migratie

- Van word 97 naar word 2000



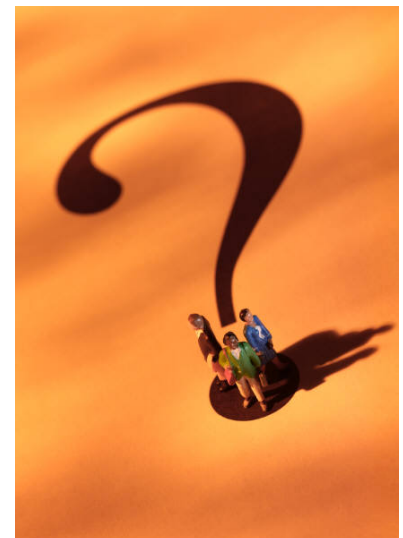
## ▪ Emulatie

- Windows op de Mac



## ▪ Standaardformaten

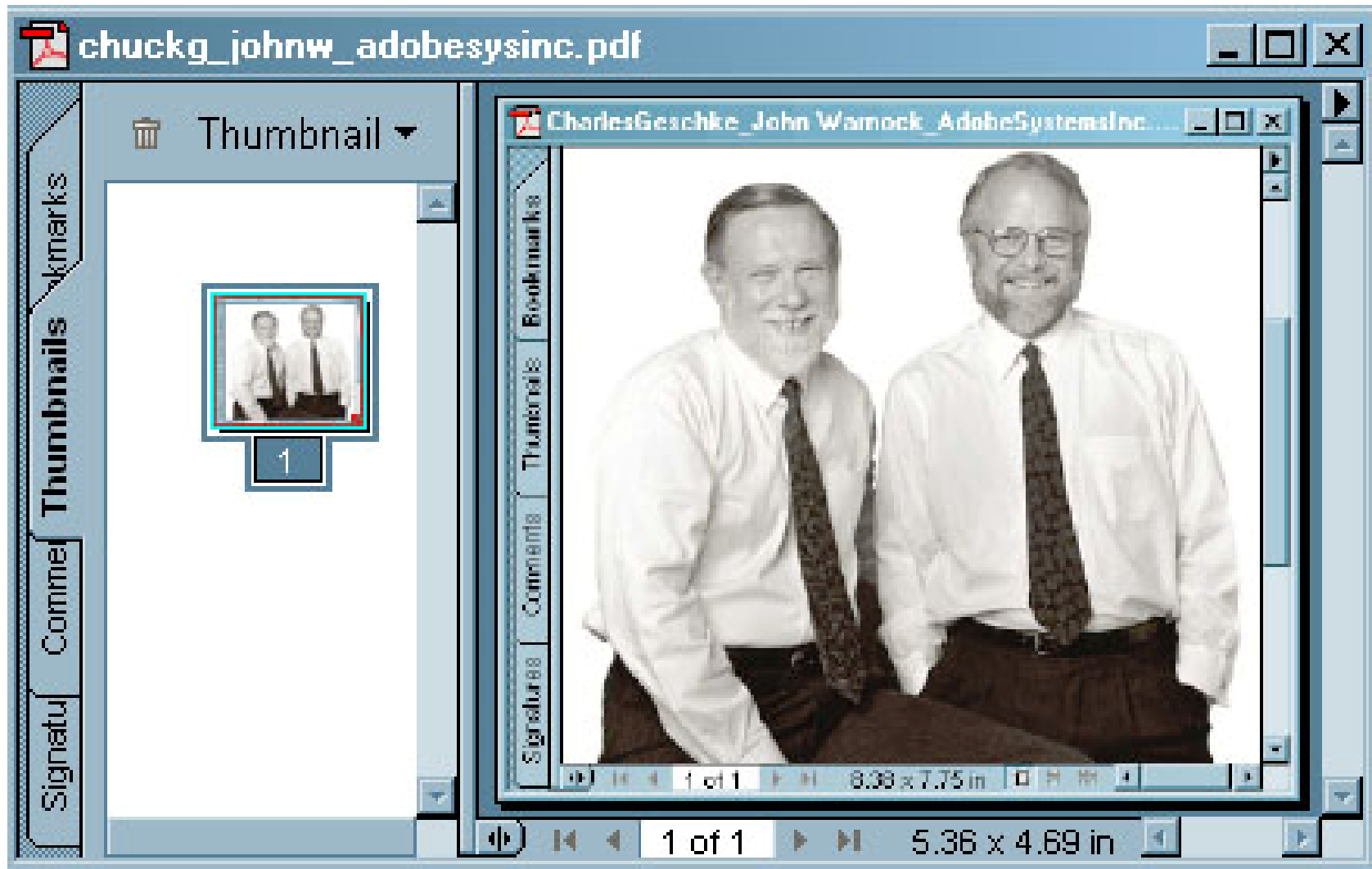
- PDF, XML, TIFF



# Achtergronden van PDF



# PDF – Op de markt gekomen in 1993



# Adobe (Acrobat) Reader was niet altijd gratis ...

## Adobe Systems Introduces Adobe Acrobat

### *Groundbreaking Software Enables Digital Document Communications Across Major Computer Environments*

New York (June 15, 1993) (NASDAQ:ADBE) — Adobe Systems Incorporated today introduced Adobe™ Acrobat™, a family of new software products that enable for the first time a practical and effective method of communicating all types of documents across different types of computers. This announcement, along with supporting announcements from technology and business partners and customers' unveiling of new products and services that utilize Acrobat software, were made today at a special event at the The Auditorium at Equitable Center in Manhattan.

#### **Pricing and Distribution**

The Acrobat Reader is priced affordably for large quantity buyers:

50 copies	\$50 per copy
100 copies	\$40 per copy
500 copies	\$35 per copy

# Een ezelsbruggetje

## Bestandsversie

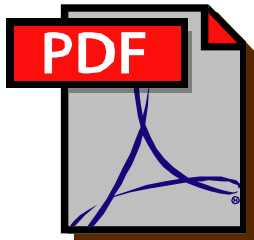
PDF 1.0  
PDF 1.1  
PDF 1.2  
PDF 1.3  
PDF 1.4  
PDF 1.5  
PDF 1.6  
PDF 1.7

## Rekensom

$1+0=1$   
 $1+1=2$   
 $1+2=3$   
 $1+3=4$   
 $1+4=5$   
 $1+5=6$   
 $1+6=7$   
 $1+7=8$

## Acrobat Versie

Acrobat 1.0  
Acrobat 2.0  
Acrobat 3.0  
Acrobat 4.0  
Acrobat 5.0  
Acrobat 6.0  
Acrobat 7.0  
Acrobat 8.0



# Adobe PDF 1.0

- Introductie in 1993
- Universele cross-platform documentuitwisseling
- Document-interne hyperlinks
- Bookmarks
- Font Embedding

The Adobe Acrobat family of software products.



Acrobat Exchange



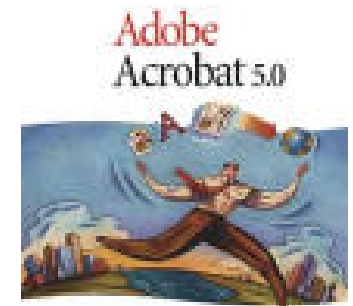
Acrobat Distiller™



Acrobat Reader

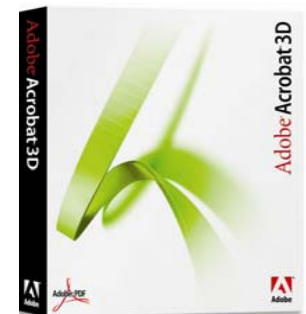
# Adobe PDF 1.4

- Introductie in 2001
- Ondersteuning van Transparantie
- Beveiliging op basis van 128-bits encryptie
- Verbeterde ondersteuning voor Javascript
- Gelabelde PDF
- PDF op ieder Device
- XML-ondersteuning



# Adobe PDF 1.6

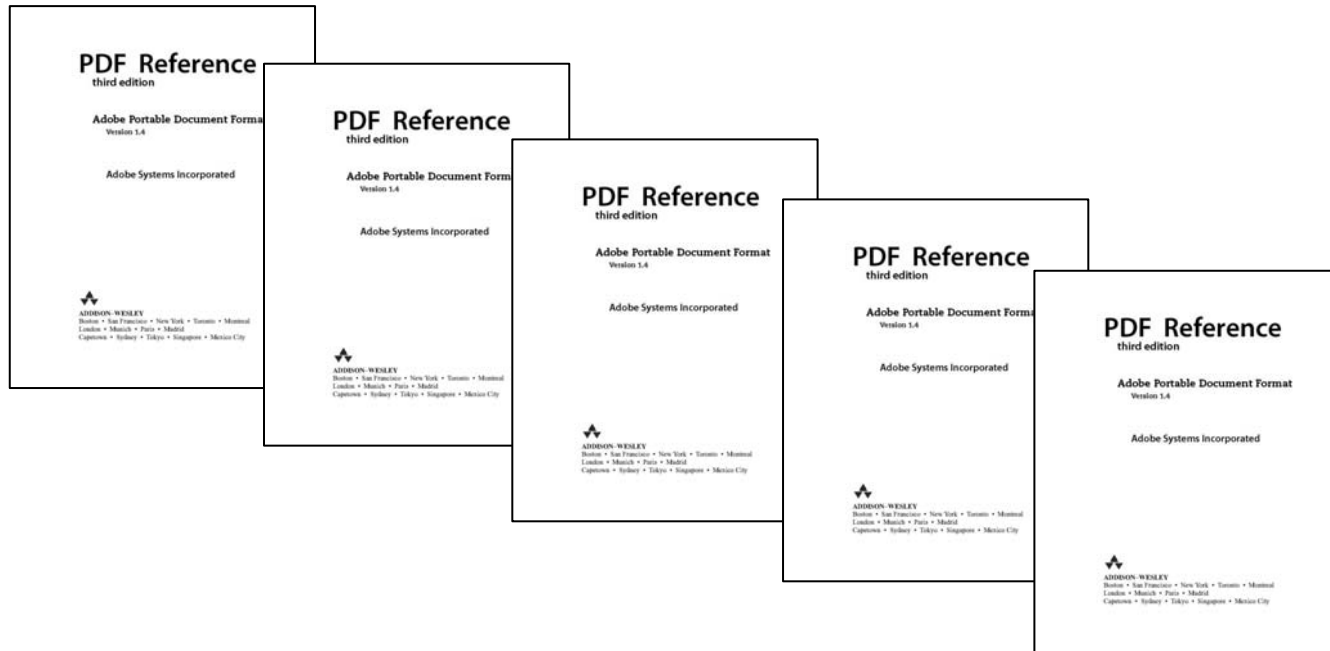
- Introductie in 2005
- Object data
- DRM mogelijkheden
- Canvasgrootte verruimd
- U3D
- Reader extensie functies
- PDF/A (en andere ISO standaarden)



# (PDF) Standaarden



# PDF heeft een open en gepubliceerde specificatie



# PDF gebaseerde standaarden

- **PDF/X—ISO 15929, 15930-1, 15930-3** 15929:
- Uitwisseling digitale advertenties
- 15930-1: Gebruik van CMYK gegevens
- 15930-3: Volledige uitwisseling van kleurenbeheer werkstromen
- PDF/x-2: Onder ontwikkeling – vervangen van beelden en XMP

- **PDF/A—NWI ISO/TC171/SC2 N226 E**
- Lange termijn bewaring van digitale documenten

- **PDF/is—IEEE—**  
Image-Streamable  
Internet Fax  
Transmission
- Fax- en internet  
standaard

- **Nog in overweging**
- PDF/Accessibility
- PDF/Engineering
- PDF/Finance

# De PDF/A Standaard



# Achtergronden van PDF/A

- Idee komt oorspronkelijk vanuit de Amerikaanse overheid
  - Administratieve diensten van de Amerikaanse rechtbanken
  - Amerikaanse overheidsparticipanten (IRS, GAO, LOC, NARA, USDA, FDA, PTO, etc.)
- Probleem
  - Kosten
  - eRecords worden een realiteit
  - Burger/klanten interactie
- **Waarom PDF als oplossing?**
  - Gepubliceerde specificatie
  - Geselecteerd als een standaard binnen veel organisaties
  - Ondersteund door een zeer actieve ontwikkelaars gemeenschap

# Waarin PDF/A verschilt van PDF

## Aanbevolen:

- Apparaat onafhankelijk formaat
- Ontwikkeling door een committee
- Derde partij oplossingen
- Gebruik van metadata
- Lossless compressie

## Vereist:

- Valide structuur labels (PDF/A-1a)
- Behoud van de visuele integriteit (PDF/A-1b)
- Inbedden van alle fonts
- Annotaties die duidelijk te indentificeren zijn

## Verboden:

- Encryptie en wachtwoordbeveiliging
- Ingebedde dynamische objecten
- Proprietary fonts of fonts die niet in te bedden zijn
- Externe hyperlinks

# Hoe maakt Adobe PDF/A

- Desktopgebaseerd
  - Adobe Acrobat Standard
  - Adobe Acrobat Professional
  - Adobe Acrobat 3D
  
- Servergebaseerd
  - Adobe Lifecycle PDF Generator ES
  - Adobe Lifecycle Forms ES
  - Adobe Lifecycle Output ES

# Waarom PDF/A belangrijk is.

- **Verlaag de kosten** van uw archiverings infrastructuur
  - Vervang bestaande documentformaten met een enkel documentformaat.
  - Open-standaarden zorgen ervoor dat meerdere platformen gebruikt kunnen worden.
  - Elimineer de overheadkosten van minder efficiënte standaarden
- **Verhoog de waarde** van uw archiefdocumenten
  - Maak gebruik van de krachtige functies van PDF om documenten te representeren en ontsluiten.
  - ISO verzekert de lange-termijn bewaring van PDF/A
- **Verklein compliancy risico's**
  - Verkrijg snel en slim toegang tot documenten ten behoeve van juridische of compliancy regels.

# Huidige participanten – gedeeltelijke lijst

- Administrative Office of the US Courts
- AFNOR
- AIIM
- ANSI
- Appligent
- BSI
- EMC/Documentum
- Glaxo Smith Kline
- Global Graphics
- Harvard University
- Hewlett Packard
- Honeywell
- IBM
- Image Solutions
- IRS
- Library of Congress
- Merck
- National Archives – US, UK, Sweden
- NPES
- PDF Sages
- Pfizer
- Victoria Archives, Australia
- Xerox

# Overheidsgebruik van PDF/A (tot nu)

## Zweden

- Nationaal archief

## Frankrijk

- Ministerie van financiën
- Ministerie van gezondheid en welzijn
- EDG GDF/GDMI (nucleaire sites)
- AFNOR

## USA

- Nationaal archief

## Nederland.....

# Huidige status en tijdlijnen

- Werk is georganiseerd door geaccrediteerde **standaard organen**
  - AIIM International (the Association for Information and Image Management)
  - NPES (The Association for Suppliers of Printing, Publishing and Converting Technologies)
- International Standards Organization (ISO) status
  - Als standaard **gepubliceerd door ISO in September 2005**
  - Werkgroep is al begonnen met het definiëren van de volgende versie( gebaseerd op de PDF 1.6 referentie)
- Ondersteuning van **verschillende leveranciers en producten:**
  - Adobe Acrobat 8, Acrobat 3D, LiveCycle PDF Generator – [www.adobe.com/](http://www.adobe.com/)
  - Visioneer - [www.visioneer.com/](http://www.visioneer.com/)
  - LuraTech - [www.luratech.com/](http://www.luratech.com/)
  - Compart Systemhaus GmbH - [www.compart.net](http://www.compart.net)
  - PDF Tools AG - [www.pdf-tools.com](http://www.pdf-tools.com)
- Meer informatie: [www.aiim.org/standards](http://www.aiim.org/standards)



✓ **PDF/A-1a**

✓ **PDF/A-1b**

# Overwegingen voor de volgende versie van PDF/A

- Gebaseerd op PDF 1.6
  - JPEG 2000 beeld compressie
  - Ondersteuning voor digitale handtekeningen
  - OpenType fonts
  - 3D
  - Audio/video content
  - Consistent met PDF/X, PDF/E, PDF/UA

# Vragen?



- [colin.van.oosterhout@adobe.com](mailto:colin.van.oosterhout@adobe.com)

**Better by Adobe.™**