

Introducción al PDF/A

Manuel Asorey
EPS



- ❖ La necesidad de preservar la documentación
- ❖ El formato PDF/A
 - ❖ ¿Qué es un PDF/A?
 - ❖ ¿Por qué PDF/A y no...?
 - ❖ Ventajas del PDF/A
- ❖ Cómo hacer PDF/A correctos
- ❖ Un vistazo hacia el futuro...



- ❖ Fundada en 1996
- ❖ Especializados en el PDF desde el primer día
 - ❖ Inicialmente en su uso en Artes Gráficas
 - ❖ Estandarización y flujos de trabajo ISO PDF/X
- ❖ Proveedores de software y soluciones PDF
- ❖ Miembros del PDF/A Competence Center





Jamdat Nasr, Uruk III; 3100–2900 A.C.



- ❖ La necesidad de preservar los documentos es tan antigua como la propia escritura
 - ❖ La escritura nace en Mesopotamia hacia el 3200 a.C.
 - ❖ El soporte utilizado es arcilla cocida
- ❖ En la actualidad nos encontramos con:
 - ❖ Ingentes cantidades de documentación en papel
 - ❖ Enormes volúmenes de información “digital”



Los problemas del papel...



- ❖ El problema más grave
 - ❖ La preservación en el tiempo
- ❖ Y además...
 - ❖ Ocupa grandes cantidades de espacio (¡con sus costes!)
 - ❖ Hace ardua cualquier búsqueda
- ❖ Solucionado parcialmente...
 - ❖ Mediante la digitalización de los archivos



Y los problemas del “digital”...



- ❖ Aún más complejos
- ❖ Problemas hardware
 - ❖ El soporte utilizado
 - ❖ El unidad lectora necesaria
- ❖ Problemas software
 - ❖ El software de creación y/o visualización
 - ❖ El sistema operativo



El formato de archivo ideal...



❖ Independientemente del soporte...

❖ Accesible

- ❖ Sin encriptación y fácil de visualizar

❖ Independiente

- ❖ De aplicaciones y S.O.

❖ Público

- ❖ A poder ser un estándar

❖ Comúnmente aceptado







- ❖ A de “Archivo”
- ❖ Estándar ISO 19005-1: 2005
 - ❖ Document Management - Electronic Document file format for long-term preservation - Part 1: Use of PDF 1.4 (PDF/A-1)
 - ❖ Nacido con la finalidad de garantizar la conservación y visualización exacta de los ficheros digitales



❖ Un PDF/A debe:

- ❖ Contener todos los elementos necesarios para su visualización fidedigna

❖ Pero no puede:

- ❖ Incluir audio, video, JavaScript, transparencias...
- ❖ Hacer referencia a contenidos externos

❖ Aunque sí puede:

- ❖ Ser firmado digitalmente y contener formularios



- ❖ Cumple con todas las características para un formato de archivo ideal
 - ❖ Es un estándar ISO
 - ❖ Es independiente de software y sistemas operativos
 - ❖ Es accesible (¡la encriptación esta prohibida!) y puede ser visualizado en cualquier aplicación PDF
 - ❖ Heredero de la amplia aceptación del PDF globalmente



¿Por qué PDF/A y no...?

- ❖ ¿Formatos ráster comúnmente aceptados en archivo como el TIFF-G4?
 - ❖ No estándar, B&N, obsoleto, pesado y con “pérdida de información”
- ❖ ¿Formatos de amplia difusión como Ms Word?
- ❖ Nuevos formatos como el XPS
 - ❖ No es un estándar ISO (si ECMA-388)
 - ❖ Limitaciones en cuanto al soporte multiplataforma
- ❖ ¿Y qué hay del propio PDF, estándar ISO 32000?



Usemos como ejemplo el folleto del PDF/A Competence Center



Escaneado en TIFF G4 (B&W): 168KB

Escaneado en JPEG: 1 MB

Escaneado y convertido a PDF/A (OCR incluido):

Desde TIFF G4: 56 KB

Desde JPEG: 136 KB

PDF/A desde fichero nativo: 68 KB



Manuel Asorey
EPS





Een waaijer van fruit en groenten

• Het is concentratie op de gezonde voedingsstoffen in een combinatie van fruit en groenten. Elk flesje van 100 ml bevat het equivalent van 100 g fruit en 100 g groenten en levert heel wat vitamines en mineralen waaronder kalium.



The quick brown fox jumps over the lazy dog
The quick brown fox jumps over the lazy dog

❖ No solo se verifica si la tipografía está incrustada...

❖ Se llega a un control glifo a glifo

❖ Gracias a la codificación Unicode



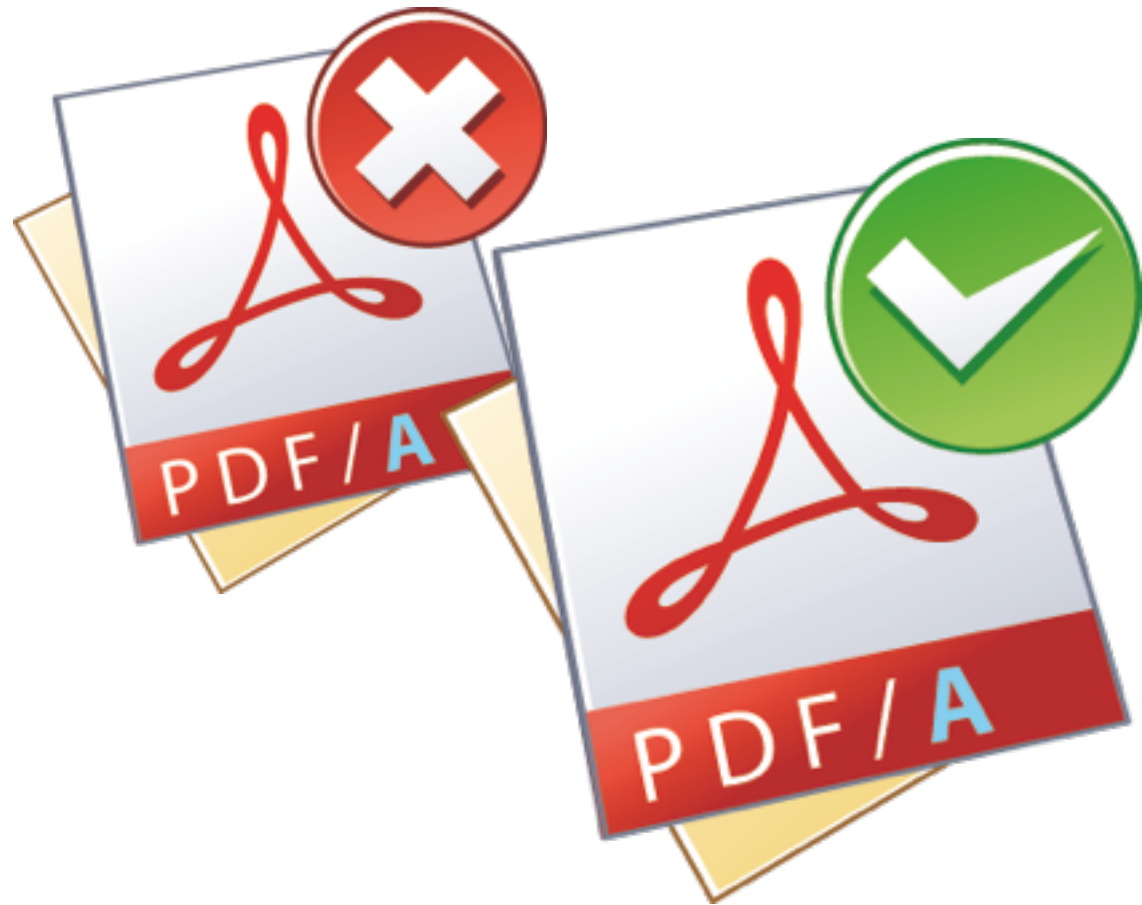


❖ Al igual que la tipografía, el color es importante...

❖ ¿Amanecía o anocheía?

❖ ¿Cuál era el logo correcto?





Tres tipos de origen posibles...

- ❖ Con origen en ficheros físicos en papel
 - ❖ Escaneo -> OCR y conversión PDF -> Validación PDF/A
- ❖ Con origen en ficheros digitales “nativos”
 - ❖ Impresión/Exportación a PDF -> Validación PDF/A
- ❖ Con origen en ficheros ya en formato PDF
 - ❖ Conversión y validación PDF/A

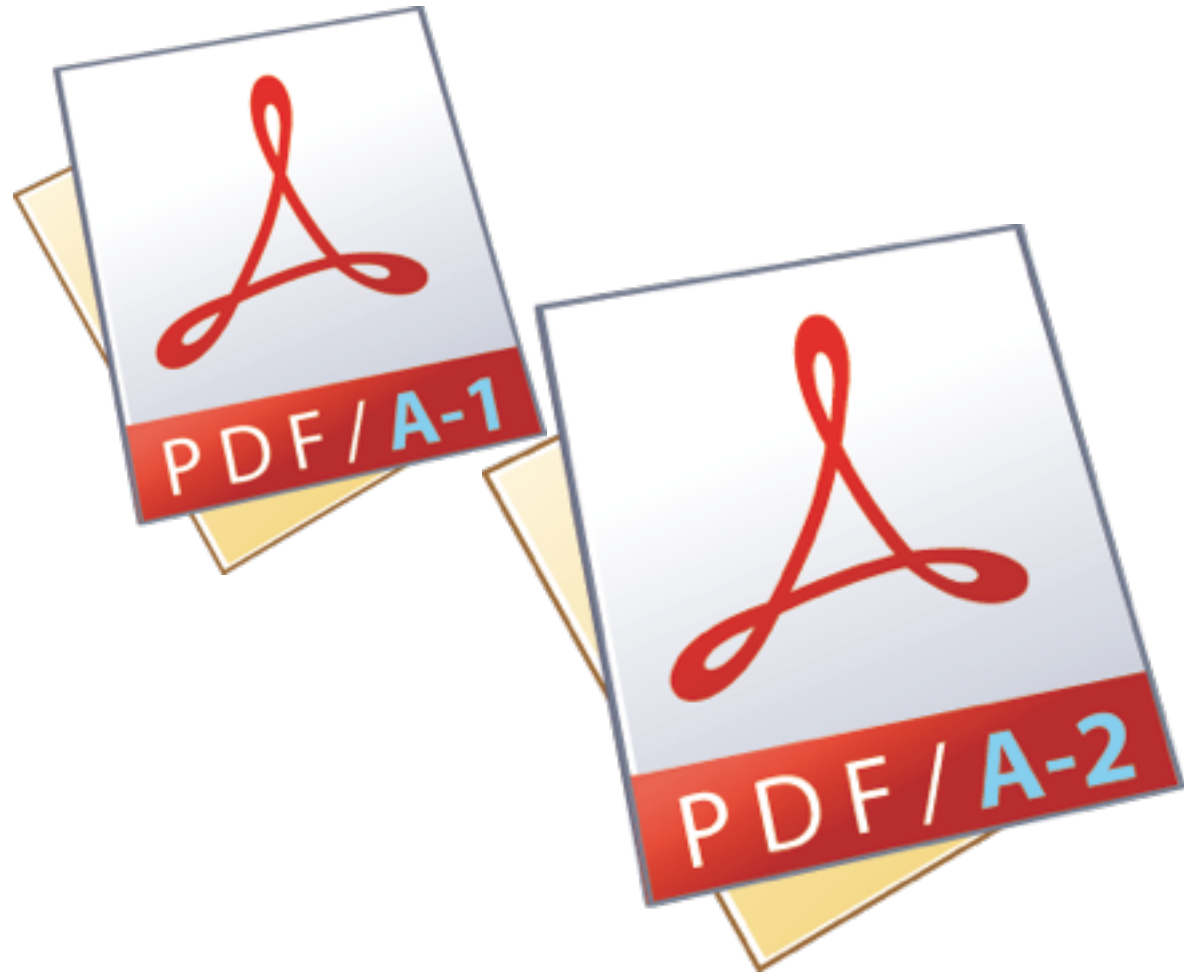


- ❖ No es PDF o PDF/A todo lo que reluce...
 - ❖ A la hora de crear...
 - ❖ Existe multitud de software que no siempre tiene la calidad deseable
 - ❖ Acrobat Reader sólo es un lector PDF: no valida
- ❖ Por tanto, es necesario un proceso de control de calidad
 - ❖ Validación PDF/A



- ❖ Existen múltiples soluciones en el mercado
 - ❖ Sólo los PDF/A validados son auténticos PDF/A
- ❖ Pero no todas son igual de fiables
 - ❖ Por ello es necesario validar al validador
 - ❖ El PDF/A CC ha desarrollado un kit de tests: Isartor
 - ❖ http://www.pdfa.org/doku.php?id=pdfa:en:isartor_test_suite
 - ❖ Complementado por el “Bavaria Report”
 - ❖ <http://www.pdflib.com/knowledge-base/pdfa/validation-report/>





Manuel Asorey
EPS



Existen dos versiones de la especificación

PDF/A-1a

“Tagged PDF” o PDF estructurado

- Permite extraer el texto ordenado

- Y la reutilización en otros dispositivos: PDA, teléfonos...

- O accesibilidad para discapacitados visuales (lectura)

PDF/A-1b

- Menos exigente que el nivel de conformidad A



- ❖ Una nueva especificación en camino: PDF/A-2

- ❖ Moderniza el estándar

- ❖ Versión de la especificación PDF actualizada

- ❖ Compresión JPEG 2000

- ❖ Transparencias...

- ❖ Que no invalida la especificación actual

- ❖ Un PDF/A-1 válido será también un PDF/A-2 válido



¡Muchas gracias por su atención!

Para más información:

Manuel Asorey
masorey@eps.es
Tel. 91 5484480



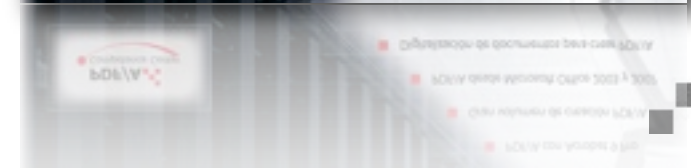
¡Muchas gracias por su atención!

A modo de agradecimiento
por su participación:

¡GRATIS!

Enviar solicitud por email a:

masorey@eps.es



¡Turno de preguntas!

