Freguntas y Respuestas frecuentes (FAQ) sobre el estándar PDF/A

PDF/A Competence Center

(Julio de 2006)

Declaracion

Este FAQ se ha preparado como apoyo del estándar ISO 19005-1:2005 (Gestión de documentos — Formato de archivo de documentos electrónicos para conservación a largo plazo, Parte 1: Uso de PDF 1.4 (PDF/A-1) ["Document management — Electronic document file format for long-term preservation — Part 1: Use of PDF 1.4 (PDF/A-1)"].

Este documento es una publicaciónde Grupo de trabajo Conjunto de PDF/A (PDF/A Joint Working Group) que trabaja bajo los auspicios del Subcomité TC-171 para Programas de Gestión de Documentos, Cuestiones de Aplicación (TC-171 Document Management Applications Subcommittee 2 Application Issues) con representantes de los Comités Técnicos 42, 46, 130 y 171 de la ISO. Se puede conseguir directamente de la ISO o de las distintas organizaciones nacionales encargadas de la estandarización (ANSI, BSI, DIN, AFNOR, AENOR, etc…).

La siguiente declaración debe acompañar a cualquier distribución de este FAQ:

> Este FAQ se puede distribuir y o traducir libremente en su totalidad. La versión autorizada vigente de este FAQ se halla en NPES y AIIM.

Qué es PDF/A

El estándar ISO 19005-1:2005 (Gestión de documentos — Formato de archivo de documentos electrónicos para conservación a largo plazo, Parte 1: Uso de PDF 1.4 (PDF/A-1) ["Document management — Electronic document file format for long-term preservation — Part 1: Use of PDF 1.4 (PDF/A-1)"] es el primero de una nueva familia de estándares ISO pensados para responder a la creciente necesidad de poder conservar la información en documentos electrónicos a

través de largos periodos de tiempo.

Como la versión inicial de PDF/A se basa en el nivel 1.4 del formato PDF, éste estándar se irá especificando por partes, de forma que las nuevas se vayan añadiendo sin que las anteriores queden obsoletas.

Así, por ejemplo, PDF/A-1 se refiere al formato definido en la Parte 1 (ISO 19005-1) del estándar (que es PDF 1.4), mientras que la Parte 2 (ISO 19005-2) y otras posteriores se podrían basar en un nivel posterior del formato PDF y podrían también definir requisitos de archivo para clases más complejas de contenido.

Porqué PDF/A

La rica naturaleza de las características del formato PDF pueden crear dificultades en la preservación de la información a largo plazo y algunas características útiles del formato PDF son incompatibles con las exigencias de la conservación a largo plazo.

Por ejemplo, los documentos PDF no necesariamente tienen un contenido autosuficiente *(not self-contained content),* sino que pueden usar fuentes del sistema y otros archivos situados fuera del documento original.

Conforme pasa el tiempo y sobre todo según avanza la tecnología, estas conexiones externas se pueden romper y las dependencias perdidas pueden hacer que se pierda información. Además, debido a la falta de estandarización de las muchas herramientas para el desarrollo de PDF que hay en el mercado, existen incoherencias en la aplicación de ese formato de documento.

Esta falta de estandarización puede ser un caos para los gestores de información en el futuro, en especial si se les hace difícil (por no decir imposible) llegar a las entrañas de los documentos PDF si no se crea una especificación de ese formato que ponga especial atención a las necesidades de conservación a largo plazo.

En todo el mundo se están creando y archivando enormes cantidades de valiosa información en formato PDF y eso hace necesario la especificación de una solución que garantice que los documentos digitales PDF seguirán siendo legibles y accesibles a largo plazo. PDF/A se ha diseñado para ser esa especificación.

Cuál es la diferencia entre PDF y PDF/A

El estándar PDF/A-1 (ISO 19005-1:2005) se basa en la referencia de formato PDF de nivel 1.4 hecha por Adobe y especifica como usar un subconjunto de componentes PDF para desarrollar programas que creen, reproduzcan y procesen una forma de PDF que sea más adecuada para la preservación de archivos a largo plazo que el simple formato PDF.

El objetivo de PDF/A-1 es preservar la aparencia visual estática de los documentos electrónicos a lo largo del tiempo, permitir el acceso en el futuro y el trasvase a otros formatos también en un futurp proporcionando el marco para:

1. La inclusión de metadatos sobre los documentos electrónicos.
2. La definición de la estructura lógica y las propiedades semánticas de los documentos electrónicos.

El resultado es un formato de fichero basado en el nivel 1.4 del formato PDF que es más adecuado para la conservación a largo plazo. Los documentos PDF/A-1 se describen mejor a si mismos, llevan en si todo el contenido necesario y son más independientes de los dispositivos de reproducción que los PDF 1.4 sin más.

Qué prohibe o permite PDF/A-1

Una de las diferencias esenciales entre el formato PDF y el estándar PDF/A es las restricciones que PDF/A impone a un PDF.

Los documentos PDF/A-1 deben incluir

- Fuentes incrustadas.
- Color reproducible independientemente.
- Metadatos XMP.

Los documentos PDF/A-1 no deben incluir

- Encriptamiento.
- Compresión LZW.
- Ficheros incrustados.
- Referencias a contenido externo.
- Transparencia PDF.
- Multimedia.
- Código JavaScript.

Qué quiere decir "largo plazo" en PDF/A

En PDF/A, "largo plazo" (long-term) se define como "el período de tiempo lo bastante largo como para que exista interés o preocupación por los impactos de las tecnologías cambiantes, incluyendo respaldo a nuevos tipos de medios y formatos de datos y a una comunidad de usuarios cambiante, para la que la información se mantiene almacenada, y que puede extenderse hacia un futuro indefinido".

¿Reemplaza PDF/A-1 a otros formatos de archivo de documentos?

No. PDF/A-1 no reemplaza otros formatos de documentos para su conservación en archivo (archival format). PDF/A-1 se ha desarrollado para permitir el uso de PDF como formato para preservación de la información de un modo sólido y bien definido.

A qué necesidades de preservación a largo plazo responde PDF/A-1

Entre los objetivos identificados de PDF/A están las siguientes características:

1. Independencia de los dispositivos. Debe poderse reproducir de forma fiable y consistente son que importe la plataforma de programa o maquinaria que se use.
2. Autoficiente en contenido. Todos los recursos y datos necesarios para su correcta reproducción se hayan dentro del documento.
3. Autodocumentado. Lleva dentro de si su propia descripción.
4. No limitado. No contiene mecanismos de protección o seguridad.
5. Abierto. Las especificaciones que lo rigen deben ser públicas y conocidas.
6. Adopción. Su uso extendido será su mejor garantía contra riesgos de preservación.

Cómo afronta PDF/A-1 estas necesidades de preservación a largo plazo

- **Independencia de los dispositivos**

  PDF/A-1 impone que los componentes del un documento sean independientes de los dispotivos, de modo que su apariencia visual se reproduzca e imprima de forma fiable y constante sin importar la maquinaria o programas que se usen. La cláusula sobre gráficos, por ejemplo, incorpora requisitos que garantizan una representación del color predecible. PDF/A-1 también prohibe el uso de componentes que no estén definidos en las especificaciones del nivel 1.4 del formato PDF.

- **Autoficiente en contenido**

  Cualquier cosa que sea necesaria para reproducir o imprimir un documento PDF/A-1 debe estar dentro del documento. La cláusula sobre las fuentes obliga a que todas las fuentes usadas para reproducir el contenido vayan incrustadas en el documento.

  Un programa creador de ficheros PDF/A-1 debe incrustar las fuentes siempre, y un programa lector que cumpla el estándar PDF/A-1 debe usar siempre las fuentes incrustadas. Esto quiere decir que el fichero se reproducirá siempre usando las fuentes que se especificaron y no las fuentes residentes en el aparato de turno.

- **Autodocumentado**

  PDF/A-1 obliga al uso de Metadatos incrustados en formato Adobe Extensible Metadata Platform (XMP). Para permitir flexibilidad en la aplicación de este aspecto, PDF/A-1 incluye recomendaciones para documentar atributos como la identificación del documento, su origen, o los metadatos de fuentes tipográficas. Además, permite que se incluyan esquemas xml que no sean XMP siempre y cuando vayan incrustados. Los desarrolladores de soluciones PDF/A-1 pueden usar XMP de muchos modos para incluir la información necesaria sobre los registros electrónicos dentro del mismo documento.

  La incrustación de los metadatos en el fichero puede aumentar el valor informativo de los documentos electrónicos y mejorar la compresión por parte de un investigador futuro.

- **No limitado**

  PDF/A-1 prohibe el encriptamiento. Esta prohibición quiere decir que no hacen falta claves o identificaciones de usuario para hacer nada con un documento PDF/A-1. Los documentos PDF/A-1 deben estar abiertos y disponibles a cualquiera con cualquier programa capaz de procesarlo. Los fabricantes de soluciones para PDF/A-1 deben proporcionar cualquier forma de control de manera externa al formato del documento.

- **Abierto**

  PDF/A-1 se basa en una especificación canónica es públicamente conocida y disponible. Cualquiera puede usar las referencias PDF 1.4 y XMP junto con las de PDF/A-1 para crear programas que puedan leer, escribir y procesar documentos PDF/A-1.

  Adobe a concedido una licencia general libre de pago de derechos *(general royalty free license)* para algunas de sus patentes para la creación de programas capaces de procesar documentos PDF/A-1. Además, Adobe a concedido a AIIM y NPES el derecho de publicar estas especificaciones en sus respectivos sitios por un tiempo indefinido.

- **Adopción**

  PDF/A-1 se ha diseñado de forma que se ponga en práctica con flexibilidad para así promover su adopción lo más amplia posible. Si esa amplia adopción ocurre, las herramientas informáticas para PDF/A proliferarán y el mercado apoyará ese formato de fichero mientrás exista su necesidad. El apoyo del mercado a PDF/A asegurará la viabilidad de PDF/A y ampliará el tiempo en el que los documentos PDF se podrán conservar como PDF/A.

Cuáles son las consideraciones respecto a las fuentes tipográficas

Muchas fuentes tipográficas tienen restricciones de uso, incrustación o intercambio. PDF/A obliga a la incrustación de las fuentes en el PDF. Por ello, las organizaciones que usen PDF/A-1 deben de tener especial cuidado en usar

fuentes tengan una licencia adecuada que permita esa incrustación.

Cuándo se debe usar PDF/A

PDF/A se debe usar como un método de estandarizar el uso de PDF para el almacenamiento de documentos electrónicos y para garantizar que estos documentos estarán disponibles en un futuro a largo plazo. Es importante atender las necesidades comerciales que requieren una reproducción fiable de los documentos electrónicos a largo plazo.

Como especificación de formato de fichero, los usuarios tendrán que establecer su propia metodología de captura para respetar sus procedimientos y políticas específicas con respecto a cosas como fiabilidad, integridad, grado de cumplimiento, alcance, etc…

Así, por ejemplo para registros permanentes en PDF, las organizaciones de la administración federal de Estados Unidos deberán usar PDF/A-1 además de los requisitos adicionales identificados en las indicaciones de la Administración de Archivos y Registros Nacional de Estados Unidos (NARA) para el traspaso de registros permanentes en forma de PDF a NARA.

Es esencial tener en cuenta que:

- PDF/A-1 por si solo no garantiza la preservación.
- PDF/A-1 por si solo no garantiza la replica exacta del material original.

¿Cuántos niveles de cumplimiento de PDF/A-1 existen?

PDF/A-1 admite dos niveles de cumplimiento. La idea es promover la creación de documentos PDF/A-1 con buena información estructural y semántica, permitiendo al mismo tiempo ficheros mucho menos complejos como imágenes escaneadas. Ambos niveles de cumplimiento se denominan Nivel A y Nivel B, respectivamente:

- **PDF/A-1a**

  En este nivel, se deben cumplir todos los requisitos de la referencia PDF conforme a las modificaciones introducidas por las especificaciones 190005 de la ISO. Las propiedades estructurales y semánticas se deben conservar.

En el nivel PDF/A-1a se usa PDF etiquetado *(tagged)* y mapas de carácter Unicode para preservar la estructura lógica del documento y el flujo de lectura del texto del documento en su orden de lectura natural. En algunos casos, para poder hacer uso de los metadatos incorporados, los usuarios necesitarán reproductores de PDF que cumplan las normas PDF/A-1.

- **PDF/A-1b**

  En este nivel se pretende que los requisitos sean los mínimos necesarios para garantizar la apariencia visual de los documentos electrónicos. En el nivel PDF/A-1b se preserva del documento original aquella estructura que permite al usuario que acceda al documento verlo y manipularlo como si estuvieran haciendolo con el mismo original.

¿Habrá revisiones o nuevas partes de PDF/A?

Ya se ha comenzado a trabajar en la Parte 2 de PDF/A con la idea de basarla en el nivel 1.6 del formato PDF. Algunos de los que van a tener que ponerla en práctica han pedido que se considere la inclusión en las partes futuras de las siguientes características:

- Compresión de imágenes con sistema JPEG2000.
- Admisión de firmado digital más avanzado y complejo.
- Fuentes OpenType.
- Gráficos tridimensionales (3D).
- Contenido sonoro (audio) o en movimiento (vídeo).
- Coherencia con otros estándares basados en PDF.

Esas nuevas partes, ¿serán compatibles con las anteriores?

La idea es que las partes futuras de PDF/A se escriban de modo que los documentos PDF/A más antiguos cumplan las partes más nuevas del estándar. La decisión final de qué se incluye en las partes más nuevas se decide en última instancia mediante la votación de los miembros de la ISO con derecho a ello.

Es esencial comprender que las partes existentes de PDF/A seguirán existiendo aunque se creen partes suplementarias, por lo que se podrá reproducir los

archivos archivados hasta un futuro no definido.

¿Admite PDF/A-1 las firmas digitales?

Sí. PDF/A-1 permite el uso de firmas digitales incrustadas en el PDF según se define en la documentación de referencia del formato PDF. Las firmas digitales en un PDF conllevan una apariencia visual, que se debe ajustar a los requisitos marcados por PDF/A-1, lo que incluyen la incrustación de fuentes y el uso de color independiente de los dispositivos.

No todas las herramientas comerciales de firmado digital cumplen estos requisitos.

Qué organizaciones van a aceptar PDF/A

Muchas organizaciones en todo el mundo están evaluando las ventajas de adoptar PDF/A y proporcionarán información sobre su criterio al respecto. Algunas de ellas ya han declarado su intención de aceptar PDF/A. Entre ellas están:

La Administración Nacional de Archivos y Registros de Estados Unidos (US National Archives and Records Administration (NARA)): NARA aceptará la transferencia de registros permanentes en formato PDF/A que adicionalmente cumplan los actuales requisitos de transferencia de registros electrónicos en formato PDF.

Los Archivos Nacionales Suecos: (Riksarkivet) Aceptarán el estándar PDF/A como una forma apropiada y aprobada de formato de archivo para algunos tipos de documentos. Esto se anunciará con una regulación técnica y de formato específica emitida por los Archivos Nacionales Suecos, que es de obligado cumplimiento para las autoridades ejecutivas suecas.

Los Archivos Nacionales Suecos no recomiendan ningún producto específico con la idea de que la independencia de productos de fabricantes concretos es una necesidad para alcanzar una preservación estable y segura a largo plazo. La política de los Archivos Nacionales Suecos es referirse ante todo a los estándares internacionales reconocidos, en segundo término a otros estándares (ya sean suecos o de otro nivel) y en último término a las especificaciones concretas que preparen los propio Archivos Nacionales Suecos.

PDF/A-1 permite la conservación de ciertas clases de documentos en u forma original. PDF/A-1 no se debe usar primariamente como formato para producir o

editar documentos, sino como formato en el que consolidar documentos. Uno de los principales criterios al establecer normas sobre formatos de documentos es la independencia con respecto a programas o maquinaria específica.

Algunos de los fabricantes que afirman que sus productos admiten PDF/A-1

Adobe, Apago, Visioneer, Callas, Compart Systemhaus, PDF Tools AG.

Porqué lleva tanto tiempo desarrollar estándares

El desarrollo de un estándar de acuerdo con los organismos de estandarización nacionales y con la política de la ISO garantiza que se hace en un proceso abierto y basado en el consenso. Varios participantes con experiencia en el área se juntan para contribuir al desarrollo.

Para conseguir que todos las partes interesadas tienen oportunidades suficientes de aportar iniciativas y de expresar sus puntos de vista, se da un proceso de votaciones muy formalizado y determinado por la organización estandarizadora.

Cómo empezó PDF/A y quién participa en su desarrollo

Las actividades sobre PDF/A se iniciaron en los Estados Unidos con el patrocinio conjunto de la AIIM, NPES y la Asociación de proveedores de tecnologías de impresión, publicación y conversión. Con los auspicios del Subcomité TC-171 para Programas de Gestión de Documentos, Cuestiones de Aplicación (TC-171 Document Management Applications Subcommittee 2 Application Issues), se formó un un Grupo de Trabajo Conjunto (WG5) con representantes de los Comités Técnicos de la ISO 42, 46, 130 y 171. Un amplio grupo de bibliotecarios, documentalistas, desarrolladores de programas para PDF, organismos gubernamentales, expertos en imagen digital y gráficos y de otros ámbitos colaboraron en el desarrollo de PDF/A.

Los primeros encuentros se celebraron a mediados de 2002 y el estándar se aprobó en junio de 2005. Expertos técnicos de quince organizaciones nacionales de estandarización participaron en todo su proceso de desarrollo.

Entre las organizaciones que actualmente participan están: AIIM, NPES, Adobe, Appligent, Callas, EMC/Documentum, Global Graphics, La Universidad de Harvard, Merck & Co, PDF Sages, la Oficina Administrativa de los Tribunales Estadounidenses, El Servicio de Ingresos Internos de EEUU y la Administración de Archivos y Registros Estadounidense.

Entre los paises que participan en estas actividades de la ISO están: Australia, China, Alemania, Francia, Japón, Holanda, Suiza, el Reino Unido, Los Estados Unidos y Suecia.

Dónde se puede conseguir más información

Hay gran cantidad de información disponible en el sitio web de AIIM. Allí se encuentran los requisitos iniciales, las actas de los encuentros, notas sobre los mismos y distintos materiales sobre PDF/A. Además, AIIM y NPES han publicado las referencias sobre el nivel 1.4 de PDF y los metadatos XMP.

Dónde se puede comprar un ejemplar del estándar PDF/A

El estándar PDF/A-1 ya publicado se puede comprar directamente a ia ISO o a cualquiera de las organizaciones nacionales de estandarización de todo el mundo.

Cómo se puede participar en su desarrollo

AIIM y NPES aceptan con gusto a representantes de fabricantes, organizaciones de usuarios y usuarios individuales. Si cree que puede contribuir al desarrollo de este estándar, no dude en ponerse en contacto con AIIM o NPES.