



Tabla de Contenido

1. GENERALIDADES

- 1.1 Objetivo
- 1.2 Alcance
- 1.3 ¿Qué es el PDF/A?
- 1.4 Breve Historia del PDF/A
 - 1.4.1 PDF/A – 1
 - 1.4.2 PDF/A -2
 - 1.4.3 PDF/A – 3

2. CARACTERÍSTICAS TÉCNICAS DEL FORMATO PDF/A

- 2.1 . Portabilidad
- 2.2 . Facilidad de creación
- 2.3 . Conversión desde cualquier fuente
- 2.4 . Seguridad
- 2.5 . Autenticidad
- 2.6 . Accesibilidad
- 2.7 . Formato no propietario
- 2.8 . Flexibilidad
- 2.9 . Facilidad de búsquedas
- 2.10 Uso de metadatos estructurados



3. EL PDF/A COMO NORMA ISO PARA ARCHIVADO A LARGO PLAZO

3.1 Versiones del PDF/A

3.2. Validación: ¿Es un verdadero PDF/A?

4 USO DEL PDF/A EN LA GESTIÓN DOCUMENTAL

4.1. PDF/A EN LA ADMINISTRACIÓN PÚBLICA

5 LOS ASPECTOS TÉCNICOS DE LA NORMA PARA LA CREACIÓN DEL PDF/A

6 OTROS USOS TÍPICOS DE PDF/A

6.2 Envío de correo

6.3 Diseño Técnico (CAD)

6.4 Trabajos Científicos

6.5 Conversión de Archivos

7 HERRAMIENTAS DE CREACIÓN Y VERIFICACIÓN DEL PDF/A

8 BIBLIOGRAFÍA

1. GENERALIDADES

1.1 Objetivo

El formato digital PDF/A es reconocido por tener características que garantizan la Preservación Digital a Largo Plazo, fundamentadas en norma técnica ISO, razón por la cual, la presente guía pretende dar a conocer y facilitar el correcto uso de este formato con miras a garantizar el acceso y la integridad de los documentos electrónicos que, por su contenido informativo, se prevea que deben permanecer disponibles, por espacios de tiempo superiores a cinco (5) años.

1.2 Alcance

Este documento está dirigido especialmente a las entidades públicas colombianas y las privadas que cumplen funciones públicas como sujetos obligados de la ley 594 de 2000.

1.3 ¿Qué es el PDF/A?

Es un formato para el almacenamiento de documentos electrónicos que asegura la permanencia de estos a largo plazo.

El PDF/A hace parte del conjunto de formatos de la familia PDF y fue diseñado exclusivamente con características orientadas al archivado a largo plazo. La Norma ISO 19005-1:2005 incluye una serie de requisitos a los programas para la visualización de archivos PDF/A, orientados a facilitar la gestión de color, el uso de fuentes integradas, o la posibilidad de realizar anotaciones por parte del usuario.

El estándar PDF/A tiene como objetivo la creación de documentos PDF/A cuya apariencia visual permanezca igual en el transcurso del tiempo. Estos resultados deben ser independientes del software y de los sistemas utilizados para crearlos, almacenarlos o reproducirlos.

1.4 Breve historia del PDF/A

1.4.1 PDF /A - 1

Es la primera versión del PDF/A, fue publicada en el año 2005 (*Electronic Document File Format For Long Term Preservation Part 1*). Esta versión está construida alrededor de la especificación 1.4 del formato PDF y es la que actualmente tiene mayor uso.

Esta versión tiene dos niveles de conformidad la de accesibilidad y la básica.

El nivel de conformidad básico (PDF /A - 1 b) es relativamente fácil de conseguir pues su foco está en la integridad visual, lo que se logra con la incrustación de las fuentes originales y el cumplimiento de las restricciones ya mencionadas.

El nivel de conformidad de accesibilidad (PDF /A -1 a) es más complejo pues requiere que se cumpla con el nivel de conformidad básico y que el PDF tenga una estructura claramente definida a fin de poder cumplir con los requerimientos de accesibilidad. Se usa un PDF etiquetado (*tagged*) y mapas de caracteres UNICODE para preservar la lectura lógica del documento y el flujo de lectura en su orden natural. Los archivos que tienen este nivel de conformidad en un futuro pueden ser leídos por un lector de e-Books.

1.4.2 PDF /A - 2

En el año 2011 se publicó la especificación PDF /A -2 (*Electronic Document File Format For Long Term Preservation Part 2*), luego de la cesión en el 2008 del estándar PDF por parte de ADOBE a ISO y está construida alrededor de la versión 1.7 del formato PDF.

En esta versión se actualizan las capacidades del PDF /A, se da soporte a la compresión JPEG 2000, soporte para el manejo de las transparencias y el uso de capas que pueden ser usadas para la creación de documentos multilingüe

Esta versión maneja los mismos niveles de conformidad que la versión anterior, el de accesibilidad y el básico y se incluye el nivel de conformidad U para documentos escaneados sobre los que se ha realizado un proceso de OCR, usando la tabla de caracteres de UNICODE.

1.4.3 PDF /A -3

A finales del año 2012 se publicó la tercera especificación el PDF /A - 3 (*Electronic Document File Format For Long Term Preservation Part 3*) que soporta la incrustación de archivos en otros formatos

Gracias a esta versión será posible archivar un plano en formato PDF /A junto con el archivo CAD original incrustado.

Esta versión al igual que la PDF/A 2 tiene 3 niveles de conformidad, el de accesibilidad, el Básico y el Unicode.

2. CARACTERÍSTICAS TECNICAS DEL FORMATO PDF/A

2.1 Portabilidad

Es la capacidad de este formato de incluir elementos tecnológicos que permiten validar su autenticidad, independientemente de la plataforma tecnológica en que se consulte. Esta característica es fundamental para cualquier formato de documento electrónico en su forma final. Y para el caso del PDF/A lo hace confiable.

Caso contrario las páginas web dependen de los recursos en línea, CSS, JavaScript y los caprichos de los navegadores, resoluciones, tamaños de ventana y así sucesivamente, la primera palabra en el nombre del formato de documento “**portátil**” lo dice todo. PDF/A está igualmente en Sistemas Operativos Windows, Mac OS X, Linux, UNIX, Android, iOS y cualquier otro.

Un archivo PDF/A contiene todo lo necesario para mostrar su contenido, restringiendo el uso de cualquier elemento, artefacto o función que pueda interferir con el acceso y la correcta visualización e interpretación. Esta característica suministra un mecanismo para representar documentos electrónicos manteniendo a través del tiempo la apariencia visual. El PDF/A no requiere de información externa para que el contenido se pueda ver correctamente. Todo está embebido en el propio archivo.

Muchas personas envían (o ponen para descargar) un DOCX, HTML, PowerPoint u otro archivo. Estos otros formatos, aunque igualmente fáciles de adjuntar a un correo electrónico, no son tan cómodos para compartir como el PDF/A. Los otros formatos mencionados no se ven de la misma forma cuando se abren en diferentes máquinas, o fácilmente pueden fallar en absoluto y no se pueden abrir. Puede haber fuentes rotas u otras configuraciones de aplicación o usuario que afectan la apariencia. Puede haber información indeseable tal como notas de diapositivas, metadatos o información de "cambios" que se quiera compartir. No es seguro que los receptores tengan la misma versión de PowerPoint (o lo que sea que se esté enviando).

Como cualquier otra tecnología, la compatibilidad con PDF/A varía. Algunos proveedores implementan todas las facetas de la funcionalidad completa del PDF/A, y ya existe la posibilidad de uso en los dispositivos móviles. Cuando se trata de un documento en formato PDF/A es totalmente predecible; funciona. Por eso grandes instituciones y agencias gubernamentales confían en PDF/A para comunicar información importante a los usuarios finales, los titulares de cuentas, los contribuyentes y otros.

2.2 Fáciles de creación

Para elaborar documentos electrónicos en PDF/A, los autores de contenido, de diseño y de estilo, pueden utilizar el software de su elección, con la seguridad de que el documento aparecerá invariablemente del mismo modo en que fue creado, independientemente de la computadora, sistema operativo, lector de PDF/A, estado de conexión, red o versión del software.

Hacer un PDF/A requiere generalmente un par de clics. Los archivos PDF/A se pueden producir por cualquier aplicación que pueda imprimir. Procesadores de texto, hojas de cálculo, software de CAD, escaneó de páginas, fotografías y cualquier contenido original se puede convertir en PDF/A y así ser confiablemente compartido o almacenado.

2.3 Conversión desde cualquier fuente

Una de las características del formato PDF/A es la capacidad de combinar páginas en PDF/A con las de otros documentos de varias fuentes. Adobe InDesign, Microsoft Word, páginas escaneadas, imágenes satelitales, capturas de pantalla; pueden ser fácilmente incluidas en un solo archivo PDF/A.

2.4 Seguridad

Los Archivos PDF/A, una vez creados, son vehículos ideales para el contenido que está destinado a una distribución limitada. Muchos programas de creación y manipulación de PDF/A pueden agregar protección con contraseña de 128 y 256 bits de alta calidad e incluir archivos adjuntos protegidos por contraseña.

2.5 Autenticidad

La característica de auto-contención de los archivos PDF/A, junto con otras características técnicas, sirve para utilizar certificados digitales estandarizados y firmas digitales. El formato PDF/A está en capacidad de manejar documentos jurídicamente válidos. El flujo de conformación deberá primero generar el PDF/A y luego firmarlo. Las especificaciones de firma electrónica en archivos PDF/A deberán



estar sujetas al cumplimiento de la norma ISO 19005, para lograr conformidad. Adicional a lo anterior es importante atender lo que dispone la Ley 527 de 1999, la Ley 594 de 2000 y demás disposiciones colombianas en materia de valor probatorio de los mensajes electrónicos.

2.6 Accesibilidad

Esta característica garantiza que quien requiere consultar una información que se encuentra en PDF/A tendrá siempre la facilidad de leerla, aun cuando sea una persona que tenga disminuidas sus capacidades de visión o audición, en el caso de los archivos que cuentan con una capa de texto, esta puede ser leída por un software de lectura.

Adicionalmente, el gobierno y algunas empresas, en especial las de servicios públicos y agencias internacionales, cada vez más exigen que los documentos electrónicos estén en formatos accesibles.

2.7 Formato No propietario

PDF es un estándar internacional. Ya no es propiedad de Adobe Systems. En 2008, Adobe cedió el control de la especificación PDF a ISO, *International Estándar Organization* (Organización Internacional de Normalización, y lo plasmó en la Norma conocida como ISO 32000. Como estándar ISO, es una tecnología de gestión transparente y democrática, no propietaria.

2.8 Flexibilidad

El estándar PDF/A ofrece una amplia gama de posibilidades para usuarios y documentos con necesidades de preservación distintas y/o particulares. El análisis documental de una Entidad específica, deberá determinar la versión y modelo de PDF/A que quiere aplicar a los documentos destinados a ser preservados en el largo plazo, dependiendo de las características de los documentos electrónicos de los que provienen o de la tecnología con la que fueron creados.

2.9 Facilidad de búsquedas

Buscar información en el contenido de uno o más documentos electrónicos, es una necesidad básica. El estándar PDF/A permite que de ser necesario se cree un texto buscable mediante OCR (Optical Character Recognition) para el caso de las imágenes digitales o de manera directa cuando el formato de creación es electrónico. El usuario y la Entidad tendrán la posibilidad, no solo de disponer de un formato de archivo electrónico para preservación a largo plazo, sino que podrán hacer uso gestionable del contenido informativo, para recuperación de información a texto completo.



2.10 Uso de metadatos estructurados

El formato PDF/A acepta información estructurada a través de metadatos estandarizados XMP y etiquetado de los componentes de la estructura. PDF/A soporta la inclusión de metadatos XMP para el registro y usabilidad versátil del documento, más allá de los propósitos específicos de la preservación digital a largo plazo.



3. EL PDF/A COMO NORMA ISO PARA ARCHIVADO A LARGO PLAZO

El PDF/A es un PDF diseñado para el archivado de documentos a largo plazo. Se sustenta por la Norma Técnica ISO 19005-1:2005. Para este caso largo plazo es la expresión utilizada para identificar la necesidad de preservar funcionalmente documentos electrónicos por espacios de tiempo superiores a 5 años. PDF/A es el primer formato de archivo electrónico convertido en estándar ISO, que garantiza que los documentos que se crean hoy, se podrán abrir y utilizar en el futuro.

El formato PDF (*Portable Document Format*: Formato de Documento Portable), desde sus orígenes (1993) ya albergaba un planteamiento de “universalidad”, dado que Adobe, la empresa fabricante, planteó como objetivo del diseño, un formato de documento independiente de la aplicación, plataforma o sistema operativo.

Para el caso del PDF/A, es el estándar ISO que define las reglas que debe cumplir un documento electrónico de formato PDF, para ser almacenado, recuperado e interpretado con objetivos de uso en el largo plazo. PDF/A proporciona un mecanismo para representar documentos electrónicos, preservando en el tiempo su apariencia visual y su funcionalidad, independientemente de las herramientas y sistemas con que se haya creado.

El Archivo General de la Nación Jorge Palacios Preciado, en razón de las características técnicas y teniendo en cuenta los resultados y las mejores prácticas internacionales en gestión de documentos, sugiere el uso de este formato para archivar documentos digitales de manera inalterable durante largos períodos de tiempo. No obstante lo anterior, es importante dejar en claro que esta decisión sobre los formatos y la tecnología que cada institución utilice para la gestión de los documentos electrónicos de archivo, solo es una decisión interna, la cual debe estar fundamentada en razones de tipo técnico y funcional. Sin olvidar que las escalas de tiempo involucradas en el archivado a largo plazo, pueden ser riesgosas para las capacidades de muchas de las herramientas con que hoy se crean documentos, que pueden verse muy diferentes de una plataforma a otra.

La ventaja significativa está representada en el uso de un formato homogéneo, basado en una especificación técnica de conocimiento público (abierto), que permite no solo preservar información por largos periodos de tiempo, sino que a la vez regula

y estandariza la creación de documentos con fines archivísticos, reduciendo la posibilidad de pérdida y la cuantiosa inversión que significa conversiones retrospectivas de ingentes masas de documentos electrónicos existentes en formatos distintos.

Las características del formato PDF/A se centran precisamente en la garantía que desde la creación del documento electrónico en este formato, éste podrá consultarse sin dificultad (visual y funcionalmente), dentro de muchos años en el futuro, por las tecnologías existentes en ese momento. Un PDF/A tiene por objeto hacer que todo el contenido de los archivos PDF sea reproducible y permanentemente accesible.

PDF es un formato muy versátil, dotado de múltiples capacidades funcionales, cada día en creciente aumento. Sin embargo, desde el punto de vista de la preservación, algunas de las potentes características del PDF, pueden significar un problema para el acceso a la información contenida en el documento, con lo que se hace necesario “restringir” ciertas acciones; “obligar” un procedimiento específico; o “condicionar” el uso de algunas funciones, para garantizar que el PDF pueda ser utilizado integralmente en el futuro lejano. Eso es lo que hace PDF/A: limitar, condicionar o prohibir ciertas funciones del PDF convencional, para que no representen un problema para la preservación a largo plazo.

Por ejemplo, en el caso de un documento que utiliza fuente Calibri, tamaño 12. Si al abrir el documento la fuente no existe en el computador, no es posible “dibujar” las letras en la pantalla. En este caso PDF/A obliga, entonces, a que las fuentes estén embebidas en el documento, lo que significa que no basta que el formato disponga del nombre de la fuente, sino que también debe existir la descripción de la misma, como parte de la información interna del PDF.

Ver un documento con fidelidad es muy importante. Si tenemos un texto o una imagen y la definición de los colores no puede reproducirse igual que en la plataforma en que se creó el PDF, podría no visualizarse el documento de manera fiel, incluso perdiendo elementos importantes del mismo. PDF/A condiciona que los colores que se usen no dependan del dispositivo, sino que sean independientes y estén no solo embebidos, sino definidos de acuerdo con reglas y estándares técnicos aceptados, para una adecuada interpretación.



Si un documento PDF está protegido con una clave y no disponemos de ella, no tenemos acceso al contenido. Esto significaría un problema para utilizar el documento en veinte (20) años, por ejemplo. PDF/A prohíbe documentos encriptados, precisamente para garantizar que el contenido pueda accederse y usarse a lo largo del tiempo.

Hay muchas otras opciones que son funcionalmente útiles en PDF (embeber archivos de audio y video, por ejemplo), que no son admitidas en PDF/A. PDF/A limita algunas de las muchas operaciones posibles en PDF y en otros casos lo que hace, es especificar la forma concreta en que deben realizarse esas acciones sobre el PDF, para garantizar la conformidad con un PDF/A.

Al momento de generar documentos que vayan a convertirse en PDF/A, es importante precisar algunas características (que bien podrían forzarse en la conversión), o que deberían tenerse en cuenta, en el propio momento de la creación:

- No deben utilizarse fuentes tipográficas que no puedan ser incrustadas en el documento.
- Los gráficos que acompañan al documento deben estar en un formato independiente del dispositivo, es decir que no dependan de la tarjeta de video, del monitor, o del controlador de imagen, etc.)
- No deben emplearse funciones de transparencia
- Es preferible que no haya enlaces externos que complementen la información (serán convertidos en texto)
- Los campos dinámicos se sustituyen por contenido (formularios, por ejemplo)
- Los contenidos de PDF/A son estáticos (omitir campos de fecha del sistema, por ejemplo)
- Las firmas digitales deben tener actualización incremental
- El Documento ya debe ser PDF/A para poder firmarlo
- El mecanismo de firma utilizado para firmar un PDF/A deberá ajustarse a las especificaciones de la ISO 19005

3.1 Versiones del PDF/A

La primera pregunta que generalmente surge al pensar en el uso de PDF/A es, lógicamente, cuál es la versión de la norma que debe utilizarse. Las opciones son las



siguientes y la decisión tiene como prerrequisito un análisis documental propio para cada Entidad:

- PDF/A-1a, PDF/A-1b
- PDF/A-2a, PDF/A-2b, PDF/A-2u
- PDF/A-3a, PDF/A-3b, PDF/A-3u

PDF/A-1 es la versión original de la norma. No permite una lista completa de características más actuales y modernas de PDF, como transparencia, algunas formas de compresión de imágenes, y capas. Todos los recursos del documento (imágenes, gráficos, caracteres tipográficos) deben estar integrados en el propio documento PDF/A.

PDF/A-2 extendió el soporte para las características restringidas en el PDF/A-1. Esta versión permite compresión, elementos transparentes y capas. También habilitó la funcionalidad de almacenar otros archivos PDF/A dentro de un archivo PDF/A-2 (permitiendo que un PDF/A-2 se comporte como un contenedor de PDF/A's).

PDF/A-3 es la versión más reciente y presenta la manera de almacenar cualquier archivo (no solo PDF/A) dentro de un archivo PDF/A-3. Esto permite que los archivos adjuntos de un correo electrónico, por ejemplo, se almacenen en su forma nativa como parte del PDF/A-3. Igualmente, facturas electrónicas convertidas a PDF/A-3, pueden guardar como parte del archivo, los datos numéricos de la factura, incrustados como XML.

NOTA sobre PDF/A-3: Un detalle de análisis para esta versión de PDF/A (PDF/A-3) obliga pensar si no se vulnera el objetivo de preservación a largo plazo previsto para el PDF/A, toda vez que un PDF/A, preparado para disponibilidad en el futuro lejano, puede contener como parte de su información, un fichero asociado e integrado, el cual no está, ni goza de la misma preparación, con lo que podría estar en riesgo la integridad y completitud de la Información.

Los “sabores” del PDF/A entregan propiedades y características específicas a cada versión, condensadas en la siguiente explicación:

La “b” de “básico”, se centra enteramente en la producción visual del documento. Lo único importante es poder ver el documento en la pantalla o imprimirlo exactamente como se creó. La “a” de “avanzado” añade unos requisitos adicionales. El texto debe estar integrado de tal manera que pueda ser extraído fácilmente (el significado debe ser claro, no solo el aspecto visual), todos los componentes del PDF/A deben ser



etiquetados (para poder distinguir elementos como encabezado, cuerpo de texto, párrafos, tablas, etc.). La “u” hace referencia a “unicode” que expande el nivel de conformidad “b” para especificar que todo el texto se puede asignar a códigos de caracteres Unicode estándar.

Con este abanico de posibilidades, podemos indicar que el uso de una versión y “sabor” de PDF/A depende directamente de los tipos de documentos que se quieren archivar y de las características que el proyecto de preservación digital desea conservar en el archivo de PDF. Los documentos en papel escaneados se pueden convertir en PDF/A, conformidad “b”, sin ningún trabajo adicional. Normalmente, el nivel “a” solo puede satisfacerse mediante la conversión de documentos nacidos digitales. Lo importante a tener en cuenta es que resulta más fácil crear un archivo PDF/A “básico”, que un archivo PDF/A “avanzado”, debido a todos los requisitos adicionales que exige contener, para que sea conforme al estándar.

3.2. Validación: ¿Es un verdadero PDF/A?

Dado que los documentos que se pretenden preservar a través del uso y aplicación del formato PDF/A, son evidencias que reflejan el actuar administrativo y sirven como prueba o investigación retrospectiva (histórica) por un espacio de tiempo muy largo, es más que necesario que se verifique que aquello que se está archivando, es realmente un PDF/A. Solo con ver un documento a través de un software lector de PDF, no es posible saber si cumple con las características de PDF/A. Muchas aplicaciones genéricas afirman el cumplimiento de la norma, pero esto siempre debe interpretarse como un indicador y no debe usarse en lugar de un examen completo de validación de conformidad. Por eso es importante validar el PDF/A, porque al tratarse de un formato computacional con especificaciones muy precisas y condiciones y reglas muy especiales, la simple interpretación de un software genérico, no es conveniente para que se pueda garantizar la conformidad necesaria con el estándar ISO 19005, y será solo a través de una revisión completa de todos y cada uno de los componentes relevantes del documento, que se certifique.

Este apartado de la Guía, entonces, se llama la atención sobre la necesidad de disponer de herramientas de software especialmente diseñadas para convalidar que el archivo PDF en el que se almacena la información, es realmente conforme a alguna de las variantes PDF/A descritas anteriormente. Es un mecanismo para asegurar que la evidencia documental cumple con los requisitos de la norma y la



Entidad garantiza de esta manera que sus esfuerzos y recursos están enfocados en una adecuada preservación de documentos electrónicos.

¿Cuándo se necesita la validación?

- ¡En todos los casos! Si lo que estamos guardando son documentos electrónicos válidos, íntegros y completos, la Entidad deberá tener la certeza que todos los PDFs que se están guardando son apropiados para el archivado a largo plazo y están conformes con la ISO 19005. En otras palabras, se debe asegurar que los documentos PDF se visualizarán y funcionarán correctamente en el futuro, a través precisamente de su validación de conformidad.
- En caso de grandes volumen documentales electrónicos institucionales, se recomienda una validación automatizada, preferiblemente integrada en una plataforma de gestión de documento electrónico de archivo, para que de forma transparente y sin esfuerzo, la Entidad logre los objetivos de preservación a largo plazo, tan importantes para la gestión archivística de los documentos.

Hay momentos en el ciclo vital de creación, recepción o uso de los PDF/As, que la validación es adecuada para determinar la conformidad con el estándar. Un primer momento es cuando se recibe un PDF. Ese puede ser el instante preciso para validar si cumple o no con la norma, porque finalmente, es imposible saber cómo se creó el archivo PDF/A. Otro momento importante es cuando el PDF se crea. Cuando la creación de PDF/A es controlada por la Entidad, ese deberá ser el momento de su validación, para asegurar que el proceso se llevó a cabo con éxito. Recibido o creado el documento, si no se ha validado, deberíamos hacerlo antes de distribuirlo. Se recomienda una validación por adelantado, antes de entregarlo o hacerlo disponible en línea, para asegurar que los documentos entregados por la Entidad cumplen con los requisitos de archivado.

Al cierre de los procesos y trámites, se deben conservar todos los atributos originales de los documentos y solo cuando se cierren éstos, proceder con la conversión a PDF/A y su respectiva validación de conformidad con la ISO 19005 para archivado. El momento apropiado de verificar conformidad, lo determinará la Entidad, con base en juicios y criterios de análisis documental y análisis procedimental, de acuerdo con sus políticas corporativas. Esta situación debe quedar documentada en el programa de gestión documental de la Entidad.



Existe una amplia gama de opciones de software especializado en la validación de formatos PDF/A y su conformidad con la ISO 19005.

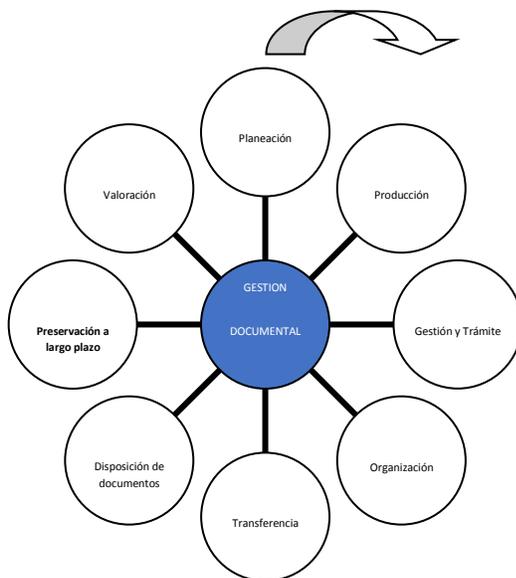
En el mercado de software con enfoque en PDF, existen desde soluciones manuales de trabajo en línea con documentos electrónicos específicos, hasta soluciones de escritorio capaces de validar cientos de miles de archivos PDF/As en muy poco tiempo. Hay opciones de costo y opciones gratuitas, para diferentes usuarios y necesidades. Las casas fabricantes de estos productos tienen a disposición de los desarrolladores, SDK (software Development Kit) o Kit de Desarrollo de Software para ser integrados en soluciones informáticas para la gestión de documentos, con lo que se automatiza la validación de conformidad. La estrategia puede ir desde una validación archivo por archivo, hasta procesos automatizados de validación y/o conversión, en los que no hay intervención humana.

Sea cual fuere la necesidad corporativa de una Entidad, la recomendación se centra en la necesidad específica de adquisición de software especializado y funcionalmente operativo para validación de PDF/A, de acuerdo con la naturaleza y volumen documental de un proyecto específico.

4. USO DEL PDF/A EN LA GESTIÓN DOCUMENTAL

La posibilidad de usar el formato PDF/A en los procesos de la Gestión Documental garantiza un flujo ordenado de actividades frente al manejo de la información allí contenida, porque se establece una metodología basada en el fácil acceso a los documentos; sin embargo el usuario tendrá diferentes alternativas en formatos para los documentos que tramita en su Entidad, pero que no podrán presentar tantas ventajas como el PDF/A.

Recordemos que los ciclos de los procesos de la Gestión Documental están representados de la siguiente manera:



Entonces es así como desde cada uno de los procesos de la Gestión Documental, se puede vincular el uso del PDF/A.

- **PLANEACION:** creando políticas, esquemas de trabajo, normas internas, manuales, instructivos y campañas de trabajo para la utilización del formato PDF/A con fines de trabajo administrativo, que más adelante se podrán volver probatorios de acuerdo a sus características.



- **PRODUCCION DOCUMENTAL:** donde se creen documentos de archivo y documentos de apoyo enfocados al cumplimiento útil de la información dentro los mayores estándares de accesibilidad, seguridad e integridad que puede brindar el formato PDF/A.
- **GESTION Y TRAMITE DE LOS DOCUMENTOS:** donde el ciclo de vida de los documentos se oriente al manejo de un formato PDF/A, desde su nacimiento hasta su disposición final.
- **ORGANIZACIÓN DOCUMENTAL:** donde los archivos electrónicos creados en formato PDF/A permitirán rutinas de trabajo más eficientes en los procesos de Ordenación, Clasificación y Descripción; esta última con la vinculación de metadatos estandarizados para toda la información que llegase a manejar la Entidad.
- **TRANSFERENCIAS DOCUMENTALES:** donde las responsabilidades de uso, trámite y difusión se adicionan desde el mismo momento en que nace el documento, por lo cual sus características embebidas hacen de su información un material único de posibilidades de trabajo.
- **DISPOSICION DE LOS DOCUMENTOS:** al permitir la interacción de la información desde un solo formato estándar, como lo es formato PDF/A, donde establecer políticas para la Conservación, Reprografía o Eliminación será mucho más fácil para la Entidad.
- **PRESERVACION A LARGO PLAZO:** para que desde la misma creación del documento se prevean acciones para garantizar la accesibilidad a la información, con total garantía de su integridad, ya que se constituye como el objetivo funcional de la presente cartilla.



- VALORACION DOCUMENTAL: donde cada persona empleada o no de una Entidad, observe en el formato PDF/A como el resultado de todo un accionar

Institucional en torno al manejo integral de la información enmarcada en un solo modelo.

4.1. PDF/A EN LA ADMINISTRACIÓN PÚBLICA

El uso del formato PDF/A dentro de la Administración Pública, resulta de importancia teniendo en cuenta que establece las condiciones adecuadas para la preservación digital a largo plazo de documentos electrónicos. Los avances que se han realizado se demuestran a través de los índices de gestión documental, tales como el FURAG, como se ve a continuación.

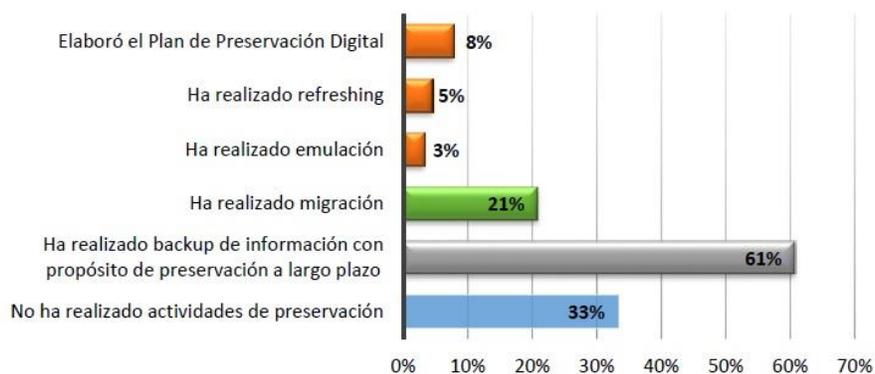
Para el año 2015, los resultados se muestran en la siguiente gráfica:



Tabla 34 .Frente a la preservación digital a largo plazo de documentos digitales y/o electrónicos de archivo, la Entidad:

Respuesta	Frecuencia	Porcentaje
No ha realizado actividades de preservación	51	33%
Ha realizado backup de información con propósito de preservación a largo plazo	93	61%
Ha realizado migración	32	21%
Ha realizado emulación	5	3%
Ha realizado refreshing	7	5%
Elaboró el Plan de Preservación Digital	12	8%

Gráfica 34 .Frente a la preservación digital a largo plazo de documentos digitales y/o electrónicos de archivo, la Entidad:



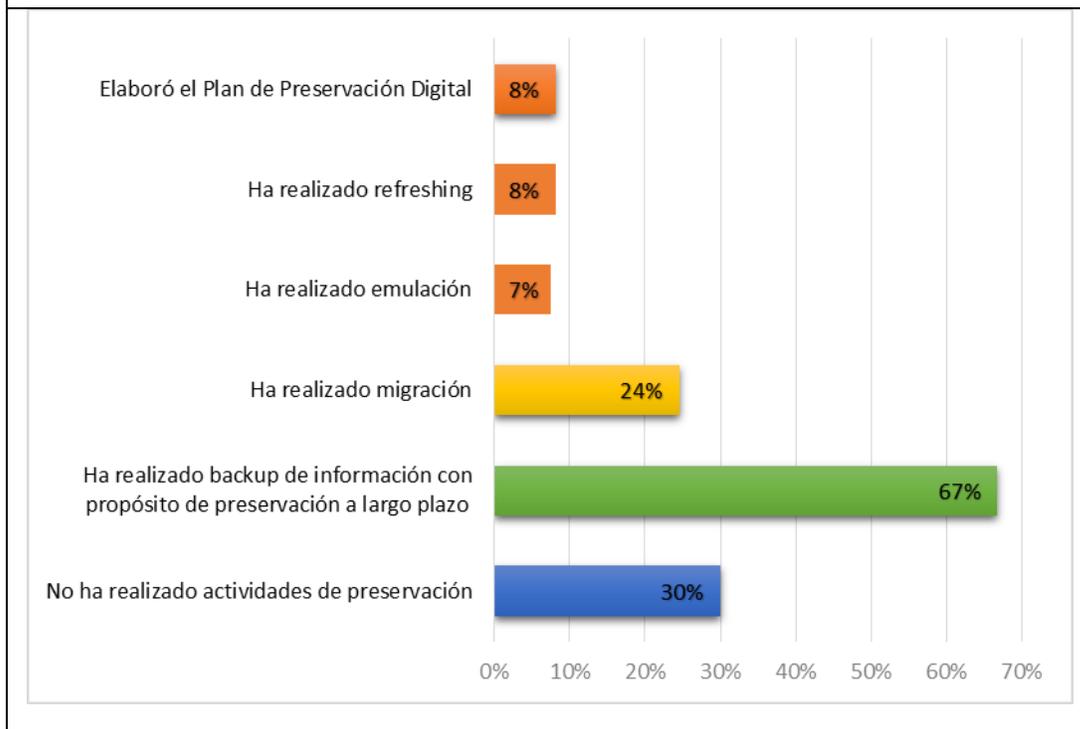
En el caso del año 2016, los resultados fueron los siguientes:

Tabla. Frente a la preservación digital a largo plazo de documentos digitales y/o electrónicos de archivo, la Entidad



Respuesta	Frecuencia	Porcentaje
No ha realizado actividades de preservación	44	30%
Ha realizado backup de información con propósito de preservación a largo plazo	98	67%
Ha realizado migración	36	24%
Ha realizado emulación	11	7%
Ha realizado refreshing	12	8%
Elaboró el Plan de Preservación Digital	12	8%

Grafica. Frente a la preservación digital a largo plazo de documentos digitales y/o electrónicos de archivo, la Entidad



Una de las principales dificultades para conocer el impacto de una estrategia en el Estado Colombiano, es la falta de información precisa y segura con respecto a

las decisiones internas en cada entidad pública, no obstante, en el año 2012, el Gobierno Nacional Expedió el Decreto 2482, orientado a establecer los lineamientos generales para la integración de la planeación y la gestión en las entidades pertenecientes a la Rama Ejecutiva. A partir de esta disposición, el Departamento Administrativo de la Función Pública, desarrolló una herramienta denominada Formulario Único Reporte de Avances de la Gestión, Furag, la cual a través de una herramienta en línea permite la captura, monitoreo y evaluación de los avances sectoriales e institucionales en la implementación de las políticas de desarrollo administrativo de la vigencia anterior al reporte.

Dentro de esta herramienta un capítulo importante es el relacionado con la Gestión Documental, el cual es responsabilidad del Archivo General de la Nación, en cuanto a las preguntas formuladas dentro de la encuesta y al análisis de las respuestas, esto con el fin de conocer la realidad de las entidades públicas de la rama ejecutiva del orden nacional en materia de archivos y así proyectar las políticas de esta entidad orientadas a mejorar este campo fundamental en el desarrollo institucional.

Como puede observarse, para el año 2015 solo el 8% de las entidades contaban con un Plan de Preservación Digital a Largo Plazo, situación que para el 2016 no mejoró, aun cuando en estos indicadores no se muestra como el formato PDF/A es tenido en cuenta como herramienta para garantizar la preservación a largo plazo de documentos electrónicos, se puede inferir, que seguramente en la mayoría de los casos este formato constituirá una de las estrategias usadas. De igual manera, la publicación de documentos como el presente, ayudará a que el PDF/A se considere por parte de la administración pública colombiana, como formato ideal para la gestión a largo plazo de documentos electrónicos de archivo.

5. ASPECTOS TÉCNICOS DE LA NORMA PARA LA CREACION DE UN PDF/A

PDF es un formato digital para la representación de documentos. Los archivos PDF pueden crearse en forma nativa, desde otros formatos electrónicos o digitalizando desde otros medios. Cuando un documento electrónico deba ser archivado durante un periodo de tiempo muy largo, la garantía de que podrá ser consultado dentro de, digamos 100 años, es una preocupación importante.

Tomando en consideración lo anterior, la International Standards Organization (ISO), publicó el 01 de octubre del 2005 el 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)*.

El propósito principal de la Norma ISO 19005 es definir un formato de archivo basado en el PDF, conocido como PDF /A, que proporciona un mecanismo para la representación de documentos electrónicos de modo que conserva su apariencia visual a pesar del paso del tiempo, independientemente de las herramientas y sistemas usados para crear, almacenar o representar los archivos.

Un segundo propósito es esta parte de la norma ISO 19005 es proporcionar un marco que registre el contexto e historia de los documentos electrónicos en metadatos dentro de los archivos conformes con la norma PDF /A.

Un tercer propósito de esta norma es definir un marco para la representación de la estructura lógica y otras informaciones semánticas de los documentos electrónicos dentro de los archivos conformes a la norma PDF /A.

Para lograr su propósito se define el formato PDF como un formato de archivo con una serie de funciones:

- Obligatorias
- Prohibidas
- Restringidas o con parámetros fijados en su uso

Estas condiciones excluyen por ejemplo funciones de seguridad como uso de contraseñas y obligan a incluir las tipografías usadas en el documento.



La información dada en los capítulos A1 a A3 proporcionan una visión general, conveniente y rápida del alcance de las diferencias de los requisitos PDF/A-1 con respecto a los del PDF Reference. La declaración normativa de todos los requisitos de PDF /A -1 se encuentran en los capítulos 2 a 6.

En el capítulo 6 de la norma se establecen los requisitos Técnicos que se deben cumplir en diferentes aspectos los archivos conformes, es decir los archivos que cumplen la norma completamente:

Estructura del Archivo: los apartados 6.1.2 a 6.1.13 tratan de temas generales del formato del archivo y sobre los elementos básicos que forman la estructura general de un archivo conforme.

Gráficos: en los apartados 6.2.2 a 6.2.10 se describen las restricciones aplicadas a los lectores y archivos conformes. Estos puntos tratan sobre los temas de representación gráfica en los que no intervienen ni fuentes ni elementos interactivos. Un archivo conforme puede especificar las características de color del dispositivo en el que se piensa representar usando una intención de salida PDF / A.

Fuentes: en estos apartados 6.3.2 a 6.3.8 se establecen los requisitos para asegurar que la representación futura del contenido textual de un archivo conforme, coincide con la apariencia estática del archivo según fue creado originalmente y para permitir la recuperación de las propiedades semánticas de cada carácter del contenido textual.

Transparencia: Se prohíbe el uso de transparencias dentro de un archivo conforme. El efecto visual de gráficos parcialmente transparentes puede conseguirse usando técnicas diferentes de las usadas por las claves de transparencia de PFD Reference.

Anotaciones: De acuerdo con la norma no están permitidos los tipos de anotaciones no definidos en el PDF Reference. Adicionalmente, no están permitidos los tipos File Attachment, Sound y Movie.



Acciones: en la Norma se establecen una serie de acciones que no pueden estar incluidas en un archivo conforme, los metadatos, la estructura lógica y los formularios interactivos.}

Metadatos: Los apartados 6.7.2 a 6.7.11 especifican los requisitos que han de tener los metadatos dentro de los archivos conformes. Los metadatos son esenciales para la gestión efectiva de un archivo a lo largo de su ciclo de vida.

Un archivo depende de los metadatos para la identificación y descripción, así como para describir los temas técnicos y administrativos apropiados.

Estructura Logia: Este tema solo es aplicable a los archivos que cumplan con la conformidad de Nivel A. En el caso de los archivos de conformidad Nivel B, este tema podrá ignorarse. Los numerales 6.8.2 a 6.8.8 dan las pautas para asegurar la recuperación del contexto textual de un archivo conformado como secuencia de palabras definidas en el orden de lectura natural del idioma en que fueron escritas.

Para ampliar esta información se debe consultar la Norma ISO 19005.

6. OTROS USOS TÍPICOS DE PDF/A

Aparte del uso ya mencionado del formato PDF/A en la elaboración de documentos electrónicos de archivo y en general en la Gestión de Documentos Electrónicos, existen otras posibilidades de uso que se mencionarán aquí, para facilitar una información completa sobre el tema.

La especificación de "Mejor edición" de la Oficina de Derechos de Autor de EE.UU¹., enumera los formatos aceptables para el depósito obligatorio de publicaciones seriadas electrónicas disponibles sólo en línea, en orden de preferencia. Para las publicaciones paginadas, PDF/A aparece primero en la lista. Otras formas de PDF son aceptables, preferiblemente con texto buscable.

La Declaración de formato recomendado de la Biblioteca del Congreso (RFS, por sus siglas en inglés) incluye PDF/A como formato preferido para trabajos textuales en forma digital, publicaciones en serie electrónicas, composiciones musicales digitales y archivos de imagen/texto acompañante para audio digital. El RFS también incluye PDF/A como un formato aceptable para otras imágenes gráficas-digitales. La lista de la RFS no distingue entre versiones de PDF/A. En general, PDF/A-1 y PDF/A-2 son formatos preferidos para documentos orientados a páginas (o principalmente textuales) cuando las características visuales y de diseño son más significativas que la estructura lógica. Tenga en cuenta que, para los PDF basados en imágenes de páginas digitalizadas, las imágenes de origen se consideran el formato maestro si está disponible. Los archivos PDF creados a partir de esas imágenes pueden optimizarse para la conveniencia del acceso en lugar de la sostenibilidad.

El PDF/A puede ser utilizado para conservar todos aquellos documentos digitales destinados a la impresión, así como aquellos no digitales que deben ser digitalizados. Por esta razón, la utilización del PDF/A es tremendamente práctica en muchos ámbitos. Los siguientes ejemplos han sido tomados de diferentes sectores en los que el PDF/A ya está siendo utilizado o en los que se están desarrollando experiencias piloto.

¹ Circular 7b. Rev: 08/2010 de la Circular 7b.

6.1 Envío de correo

Un aspecto importante a tener en cuenta lo constituye el hecho de que el correo enviado en papel a menudo debe ser conservado en un formato digital fiable. Es en este punto donde el PDF/A destaca y, por ello, están disponibles en el mercado diferentes soluciones que permiten su generación a partir de la impresión.

6.2 Diseño Técnico (CAD)

Los archivos de diseño técnico habitualmente precisan de software especializado solo para poder ser visualizados. La conversión de dichos archivos a PDF/A posibilita no solo su visualización con un software estándar sino también garantiza que dichos diseños podrán ser visualizados e impresos con el paso del tiempo.

6.3 Trabajos científicos

Múltiples universidades solicitan en la actualidad que tanto tesinas como tesis sean enviadas en formato PDF/A si es posible.

6.4 Conversión de archivos

Numerosos documentalistas no están conformes con sus archivos por distintas razones. Además, nuevas disposiciones legales solicitan nuevos requisitos para el archivado de documentos. Numerosas empresas han comenzado ya la migración de sus heterogéneos y voluminosos archivos hacia el PDF/A.

7. HERRAMIENTAS DE CREACIÓN DEL PDF/A

La creación de archivos PDF/A es relativamente sencilla, y se puede hacer a través de muchas opciones y/o canales de apoyo tecnológico de modalidad gratuita; por medio de paquetes de instalación en cada equipo; en línea, a través de Internet, etc. Se encuentran soluciones que van desde creación de ficheros individuales, en línea, libre de pago; hasta mecanismos de integración en software especializado, en donde el trabajo de creación/conversión se hará, incluso sin darnos cuenta.

Las herramientas disponibles tienen, en términos generales, más o menos un mismo enfoque tecnológico de resultado, en donde los PDF/As gozan de cumplimiento de conformidad con el estándar para el que hayan sido preparadas. Así mismo, y en relación directa con la funcionalidad que busca una entidad Archivística, es su propia responsabilidad evaluar las alternativas disponibles en el mercado en relación directa y proporcional con sus necesidades, teniendo como referencia especial el volumen y la verificación automática de cumplimiento para que la herramienta seleccionada, cumpla el objetivo que se persigue.

El volumen documental y los niveles de automatismo que se requieren en los Archivos, son sin duda, de los criterios más importantes a la hora de definir la selección de una u otra alternativa tecnológica para la generación de archivos PDF/As.

Desde el punto de vista del volumen, es casi imposible pensar en alternativas “modestas” de trabajo por demanda y con número de conversiones limitadas, como las que normalmente ofrecen las herramientas gratuitas disponibles en línea. De igual forma, desde el punto de vista de la automatización de flujos documentales, es muy poco probable que una solución “manual” resulte de utilidad, toda vez que el trabajo con mucha intervención humana, a nivel de usuario, especialmente en la selección y cargue de documentos para el proceso de conversión y/o validación, puede afectar de manera significativa cualquier intención para una gestión documental profesional. No significa esto, insistimos, en que las alternativa manuales o por demanda, estén fuera de una realidad funcional válida, sino que deben ser corroboradas de acuerdo con la realidad documental del proyecto de gestión documental de cada entidad.



Con este panorama, y advirtiendo que cualquier de decisión requiere un análisis especial de caso, es posible intuir que la mejor solución es encontrar esta funcionalidad embebida en un software profesional para la gestión de documentos, por medio de un SDK especialmente diseñado para convertir y validar cada uno de los objetos PDF o PDF/A que convivan en el sistema.



8. BIBLIOGRAFÍA

- PDF/A, PDF for Long-term Preservation
<https://www.loc.gov/preservation/digital/formats/fdd/fdd000318.shtml>
- PDF/A: Formato de archivo de documentos
<http://www.biblogtecarios.es/joaquinhierro/pdfa-formato-de-archivo-de-documentos/>
- PDF/A FAQ <https://www.pdfa.org/pdfa-faq/>
- FURAG Consultado de la página <http://observatoriosna.archivogeneral.gov.co/>



ARCHIVO
GENERAL
DE LA NACIÓN
COLOMBIA

GUIA DE USO DEL PDF/A

Versión para consulta pública

**DOCUMENTO ELABORADO POR EL COMITÉ TÉCNICO DE PRESERVACION
DIGITAL PERTENECIENTE AL SISTEMA NACIONAL DE ARCHIVOS
ARCHIVO GENERAL DE LA NACION JORGE PALACIOS PRECIADO**