



OpenSoft

" Inventer, c'est penser à côté "  
Albert Einstein

Volume

14

[www.Open-soft.fr](http://www.Open-soft.fr)

Les livres blancs

# Pourquoi Le format PDF /A ?



Patrick Barantal :



[patrick.barantal@open-soft.fr](mailto:patrick.barantal@open-soft.fr)



09 50 19 51 72



N'hésitez pas à nous rejoindre sur notre site :

[www.open-soft.fr](http://www.open-soft.fr)

Version 1.0.1

# Table des matières

<b>HISTOIRE DU FORMAT PDF:</b> .....	<b>3</b>
<b>LA NECESSITE D'UN STANDARD DANS LES FORMATS D'ARCHIVAGE.</b> .....	<b>3</b>
DEFINITION DE LA DEMATERIALISATION : .....	3
LES AVANTAGES ET LES INCONVENIENTS DE LA DEMATERIALISATION : .....	4
<b>LES EVOLUTIONS DU FORMAT PDF :</b> .....	<b>5</b>
1999 VERSION 1.3 : .....	5
2001 VERSION 1.4 .....	5
2003 VERSION 1.5 .....	5
2005 VERSION 1.6 .....	5
2006 VERSION 1.7 .....	5
<b>LES DIFFERENTS FORMATS PDF :</b> .....	<b>5</b>
PDF/A .....	5
<i>Sont exclus Du PDF/A :</i> .....	6
<i>Il est ajouté au PDF/A :</i> .....	6
PDF/A-1A : LE NIVEAU LE PLUS STRICT .....	6
PDF/A-1B: SUFFISANT POUR DES PRESENTATIONS.....	6
PDF/A-2 : LE FUTUR .....	6
PDF/X – UNE NORME PDF POUR LES DONNEES DE PREPRESSED .....	7
PDF/X-1A : .....	7
<b>POUR EN SAVOIR PLUS :</b> .....	<b>8</b>

## Histoire du format PDF:

---

Depuis 2006, un format de fichier s'est imposé dans le domaine de l'archivage long terme des documents : c'est le format PDF /A.

Avant d'expliquer le format PDF /A, il est indispensable d'expliquer l'historique de ce format, ce qui permettra de comprendre comment ce format de fichier qui vous est devenu familier sans le savoir, s'est imposé comme un standard.

Le PDF acronyme de **P**ortable **D**ocument **F**ormat a été initialement créé par la société Adobe Systems, pour un usage interne, afin de gérer ses échanges de document. En effet un document créé sur une machine ne pouvait être lu sur une autre.

Le PDF est un langage de représentation de documents, très riche en fonctionnalité, il n'a cessé d'évoluer au travers de ses différentes versions (Version 9 au moment de la rédaction de ce document).

Les documents PDF peuvent être simples ou complexes, structurés ou non, contenir du texte, des graphiques des images et même des contenus multimédia (sons, images animées etc..).

C'est la gratuité de son lecteur Acrobat Reader qui a permis sa généralisation, en effet pour rendre lisible un document au format PDF un logiciel nommé lecteur ou visionneuse est nécessaire.

Cela rappelle un peu l'histoire du format HTML (1192) qui nécessite un navigateur pour rendre les fichiers du web lisible.

Les spécifications de la version 1.3 ont été publiées en 1999, lors du lancement de la version 4 d'Acrobat Reader. Celles qui ont servi de référence pour la norme PDF/A et de PDF/x sont PDF 1.4 associée à la version 5 d'Acrobat.

## La nécessité d'un standard dans les formats d'archivage.

---

La réforme du Code Civil du mois de mars 2000 a introduit la reconnaissance de l'écrit numérique dans le droit français.

La dématérialisation des documents et des procédures est devenue une réalité en France.

### Définition de la dématérialisation :

La dématérialisation est la transformation de supports d'informations matériels (souvent des documents papier) en des fichiers informatiques.

Dématérialiser en informatique, signifie transférer sur un support numérique des d'informations qui existent sous forme analogique (papier, film, microfilm). Dans ce cas précis, on parle de la numérisation d'un document. Dans un autre cas, le document est directement crée sous forme numérique et traité comme tel.

La dématérialisation peut servir à gérer de façon totalement électronique des données ou des documents qui transitent au sein des entreprises et/ou dans le cadre d'échanges avec des partenaires (administrations, clients, fournisseurs,...)

### **Les avantages et les inconvénients de la dématérialisation :**

Les avantages de la dématérialisation sont nombreux :

- Gain de temps dans la recherche d'un document
- Réduction des frais postaux : envoi par courriel électronique.
- Economie dans l'archivage des documents
- Possibilité d'appliquer des traitements automatisés.

Mais la dématérialisation n' a pas que des avantages, en effet elle pose des problèmes techniques complexes.

- Gestion de l'intégrité d'un document
- Sauvegarde et choix des Supports de stockage
- Conservation a long terme des documents.

Le choix d'un format d'archivage s'est rapidement imposé comme un casse tête, en effet il existe plus de 200 formats pour les images et chaque année voit surgir un nouveau format d'encodage sonore ou vidéo, qui n'a pas eu un problème de CODEC.

Le choix d'un standard s'est imposé

### **C'est le format PDF/A qui a été retenu comme norme ISO.**

En septembre 2005, l'Organisation internationale de normalisation (ISO) a approuvé la nouvelle norme PDF/A pour l'archivage de documents électroniques. Selon la norme ISO 19005-1.

<http://www.iso.org/>

## Les évolutions du format PDF :

---

A ce jour, le format PDF en est à la cinquième version depuis sa première publication.

### 1999 Version 1.3 :

PDF 1.3 intègre complètement le niveau III de Postscript. Il supporte la séparation des espaces de couleur et les profils ICC (International Color Consortium ISO 15076-1-2005 défini les profils techniques des numériseurs et des imprimantes)

### 2001 Version 1.4

Introduction de la transparence ainsi que les TAGS (Gestion de la structure des fichiers). Les fichiers en niveaux de gris peuvent être compressés suivant la méthode JBIG 2 (Joint Bi-level Image Experts Groupe norme ISO/IEC 14492-2001. Taille maximum des documents 5,08 mètres à échelle 1.

### 2003 Version 1.5

JPEG 2000 peut être utilisé comme moyen de compression des images fixes, il est aussi possible de gérer des couches (layers).

### 2005 Version 1.6

Cette version supporte les « Open-Type Fonts » ; la longueur des pages peut atteindre 381 Km ; il est possible de gérer des fichiers contenant des objets en 3 dimensions.

### 2006 Version 1.7

L'encodage des documents est réalisé en UNICODE (support de tous les alphabets). Plusieurs fichiers PDF peuvent être intégrés dans un seul fichier. LE 1<sup>er</sup> Juillet 2008 PDF 1.7 est devenu la norme ISO 32000.

## Les différents formats PDF :

---

### PDF/A

PDF/A est une famille de normes ISO (ISO 19005-X) qui définissent des sous ensemble du format PDF à des fins d'archivage à long terme de documents numériques.

Cette première norme PDF/A 1 est le résultat de trois années de travail de la commission 171 de l'ISO.

PDF/A-1 est basé sur la version 1.4, dont certaines fonctionnalités n'ont pas été retenues pour des raisons de portabilité.

### Sont exclus Du PDF/A :

- La transparence
- L'inclusion de séquences sonores
- Les images animées
- Les images comprimées JPEG 2000

### Il est ajouté au PDF/A :

- Incorporation des polices de caractères nécessaires à l'affichage du texte (ce qui explique qu'un fichier PDF/a peut être de taille supérieure à un fichier PDF)
- Ajout de métadonnées au format XMP.

Le PDF/A est un dérivé du PDF qui fournit un mécanisme représentant les documents électroniques d'une façon qui conserve leur aspect visuel à long terme, peu importe les outils et systèmes utilisés pour la création, le stockage ou le rendu des fichiers.

Cette méthode de conservation rend le format PDF autonome. Le PDF/A en arrive à cette autonomie en intégrant l'information (contenu, couleur, polices, images, texte, etc.) requise pour l'affichage du document dans le document lui-même. Ainsi, le PDF/A n'a pas besoin d'information extérieure additionnelle pour l'affichage. Cependant, pour que ce format garde son autonomie, il doit exclure certaines caractéristiques de fichiers PDF standard comme des films, des sons et de la transparence.

La classification PDF/A se divise en deux parties : PDF/A-1 et PDF/A-2. La première classification, le PDF/A-1 se sous-divise en deux catégories: PDF/A-1a et PDF/A-1b. La méthode d'extraction de texte constitue la principale différence entre le PDF/A-1a et le PDF/A-1b :

#### **PDF/A-1a : Le niveau le plus strict**

- Ce niveau, également connu sous le nom de conformité de niveau A, est entièrement conforme à la norme ISO 19005-1 Standard. Cette version intègre l'étiquetage de façon que le texte puisse être extrait et visualisé par plusieurs dispositifs, incluant des appareils portatifs.

#### **PDF/A-1b: suffisant pour des présentations**

- Conformité de niveau B – Cette catégorie est considérée comme le niveau de conformité minimal. Ce niveau assure l'affichage et la lecture d'un document sur un moniteur d'ordinateur, mais ne garantit pas la lisibilité de ce document.

#### **PDF/A-2 : Le futur**

- La norme ISO 19005-2 (plus connue sous PDF/A-2) basée sur la norme ISO 32000 c'est-à-dire la version 1.7 PDF est en cours de rédaction, publication prévue fin 2010.
- Ajout récent à la norme PDF/A en cours de formulation par le Comité technique. Essentiellement, l'élément PDF/A-2 ajoute aux caractéristiques les

plus récentes de la référence PDF, comme les signatures numériques. A été ajouté : Support JPEG200, Mode transparent, Calques, PDF packages plusieurs fichiers PDF dans un fichier PDF.

## **PDF/X – Une norme PDF pour les données de prépresse**

Qu'est-ce que le PDF/X?

Le format PDF/X, comme le PDF/A, est un sous-ensemble du PDF. Le PDF/X fournit aux concepteurs, aux illustrateurs, aux ingénieurs et aux artistes graphiques un format de fichier électronique que peut imprimer correctement tout fournisseur de service. Le PDF/X assure une uniformité totale même si les fichiers sont traités à de multiples emplacements par des personnes utilisant du matériel différent. Ce format est idéal pour la plupart des transferts intercompagnies de matériel prêt à imprimer dans les cas où l'expéditeur de l'information prête à imprimer et le destinataire n'ont pas un lien important. De plus, en fournissant un format de livraison solide pour les tâches d'impression, le PDF/X offre aussi d'autres avantages incluant entre autres un visualiseur de fichier, une compression supérieure (format de fichier plus petit), un support des couleurs et une méthode d'identification de l'état de l'impression pour laquelle le fichier a été préparé. Cependant, comme pour le PDF/A, les avantages du PDF/X sont annulés par les compromis nécessaires. Des caractéristiques comme la transparence, l'encryption et la compression JBIG2 ne sont pas permises dans le format PDF/X.

## **La norme PDF/X se divise en trois catégories :**

### **PDF/X-1a :**

Cette norme est idéale dans le cas de personnes qui envoient des fichiers et qui veulent garder le plus de contrôle possible sur la tâche d'impression. Ce format est adapté aux échanges aléatoires et est conforme aux exigences les plus courantes dans de nombreuses parties du monde.

### **PDF/X-3 :**

La différence la plus importante entre les formats PDF/X-3 et PDF/X-1a est le fait qu'un fichier PDF/X-3 contient des données gérées par couleur. Cela signifie que tout outil qui peut lire un PDF/X-3 devrait aussi pouvoir lire un fichier PDF/X-1a.

### **PDF/X2 :**

Contrairement aux deux premiers formats mentionnés ci-dessus, qui sont conçus pour les échanges aléatoires, le PDF/X2 est davantage conçu pour les échanges qui nécessitent plus de dialogues entre l'expéditeur et le destinataire. Il s'agit d'un surensemble du format PDF/X-3, qui est un surensemble du format PDF/X-.

Ce que cela signifie pour votre entreprise...

La norme PDF/X est conçue pour être appliquée dans différents secteurs et entreprises dans le domaine de l'imprimerie. Si vous recevez des fichiers prêts à imprimer, assurez-vous d'indiquer clairement quels types de fichiers vous pouvez accepter (PDF/X-1a, PDF/X-3 ou PDF/X2). De plus, nous vous conseillons d'utiliser les versions les plus récentes des

outils PDF/X au fur et à mesure qu'elles sont offertes. Ces versions plus récentes vous permettent de lire des fichiers plus anciens. La plupart des produits qui sont compatibles avec le format PDF de base sont également compatibles avec les fichiers PDF/X. Vérifiez tout fichier à imprimer que vous recevez afin de vous assurer qu'il est conforme avec votre version PDF/X. De plus, si vous êtes un imprimeur ou un éditeur, assurez-vous de la compatibilité au format PDF/X à toutes les étapes du travail. Si vous envoyez des données prêtes à imprimer à un éditeur ou à un imprimeur, examinez soigneusement les options s'appliquant aux sous-ensembles du PDF/X pour établir ce qui est le mieux adapté pour votre travail. En fin de compte, le PDF/X doit vous permettre d'envoyer des fichiers numériques complets en vous donnant l'assurance qu'ils passeront à la prépresse sans erreur ou modification.

## **Bientôt... PDF/E et PDF/UA**

### **PDF/E**

Le PDF/E est une future norme qui met l'accent sur le secteur de l'ingénierie. Il traite notamment certaines des caractéristiques les plus récentes du format PDF qui servent principalement dans ce secteur, comme les métadonnées du niveau d'objet et la modélisation 3D.

### **PDF/UA**

Les lettres «UA» signifient «accessibilité universelle». Cette norme, encore en cours de préparation, met l'accent sur l'accessibilité du PDF en développant les principes de la Section 508. Le PDF/UA assure la facilité d'emploi et l'accessibilité des PDF pour tous les utilisateurs, peu importe leur incapacité.

## **Pour en savoir plus :**

---

L'association Internationale ADDS (Association for the Digital Document Standards) créée en 2006 est à l'origine de l'initiative PDF/A Competence Center, dont l'objectif est de promouvoir les échanges d'information à long terme.

[www.pdfa.org](http://www.pdfa.org)