

RÉPUBLIQUE FRANÇAISE

Ministère de la culture et de la
communication

Direction générale des patrimoines

Service interministériel des Archives de
France

Note d'information DGP/SIAF/2012/011 en date du 30 août 2012 relative à l'étude sur le format des fichiers PDF

Le directeur chargé des archives de France

à

Mesdames et Messieurs les directeurs des services centraux nationaux des archives nationales
Mesdames et Messieurs les directeurs des services départementaux d'archives
sous couvert de Mesdames et Messieurs les préfets de région
et de Mesdames et Messieurs les préfets de département

Référence : Note d'information DGP/SIAF/2011/025 en date du 15 décembre 2011 relative au lancement d'une étude sur les formats PDF

Pièces jointes : Guide méthodologique. Le format de fichiers PDF, août 2012

Comme suite à la note d'information citée en référence, j'ai le plaisir de vous informer de la parution du « Guide méthodologique. Le format de fichiers PDF », août 2012, réalisé par la société NUMEN et le CINES, à la demande des Archives de France. Ce guide constitue, comme cela vous avait été annoncé, la première partie de l'étude visant à faire le point sur les différents formats PDF.

Le format PDF existe depuis 1993¹ et a été normalisé ISO pour la première fois en 2008. Durant ses 18 ans d'existence, celui-ci a évolué avec sept révisions. Plusieurs normes ISO ont par ailleurs été élaborées pour préciser comment utiliser les fichiers PDF dans des domaines spécifiques. Le guide a par conséquent pour objectif de mieux comprendre quelles sont les fonctionnalités de ces différents versions et de comprendre le lien entre ces formats et les normes ISO élaborées à partir de certaines des versions.

¹Bien que propriété de la société Adobe, le format des fichiers PDF a toujours été publié.

Le guide se décompose en plusieurs parties : la première vise à donner une vision globale des formats PDF en décrivant, après avoir expliqué les origines du format (Postscript), ses principales caractéristiques (portabilité, compression, indépendance des polices, génération de fichier en un passage, accès direct à chaque page, mise à jour incrémentale, extensibilité), en présentant la structure d'un fichier PDF et en détaillant ses fonctionnalités (couleurs, polices, transparence, préférences de document², métadonnées, sommaire, signature, articles³, présentations⁴, structure logique, contenus facultatifs, annotations, liens hypertextes, compression, sécurité, formulaires interactifs, multimédia, illustrations 3D, programmation, fichiers embarqués, OPI⁵, CAO⁶, informations géomatiques, capture de pages Web en PDF). Un tableau récapitulatif permet d'indiquer, pour chaque version, quelles sont les fonctionnalités présentes.

Une introduction aux normes basées sur le PDF est ensuite proposée, ainsi qu'un schéma des liens entre les différentes versions de PDF, ainsi qu'entre les versions de PDF et celles du logiciel Acrobat.

Cette première partie se termine par un glossaire général, un glossaire spécifique à PDF (traduction des termes anglais/français et français/anglais) ainsi qu'une bibliographie.

Dans la seconde partie, chaque version de PDF depuis la première version 1.0 jusqu'à la future version 2.0 est détaillée sous forme d'une notice signalétique⁷ spécifiant chacune des fonctionnalités listées ci-dessus (suivant leur présence ou non pour la version étudiée).

La dernière partie est consacrée à la présentation des fonctionnalités des normes PDF suivant la même grille : PDF/A (utilisé dans le cadre de l'archivage pérenne), PDF/E (utilisé pour la documentation d'ingénierie), PDF/UA (traite de l'accessibilité en prise en compte des handicaps), PDF/VT (exigences pour l'édition), PDF/X (série de normes créées pour l'échange de documents pour l'imprimerie).

Le bureau des traitements et de la conservation se tient à votre disposition pour tout renseignement complémentaire.

Le directeur, chargé des Archives de France



Hervé LEMOINE

² Ensemble de valeurs qui spécifie les paramètres d'affichage ou d'impression d'un document.

³ Un article dans PDF est un terme très précis. Il est utilisé dans le cas où un document contient plusieurs flux de texte indépendants, comme, par exemple, dans un magazine où on pourrait avoir un article principal et un deuxième flux de texte qui continue en encadré sur plusieurs pages.

⁴ Une présentation est une façon de visualiser une série d'images sous la forme d'un diaporama .

⁵ *Open Prepress Interface* : mécanisme qui permet de mettre des images de basse résolution dans un fichier pour l'alléger lors de sa création et de les remplacer par des images haute résolution lors de son impression finale.

⁶ Conception assistée par ordinateur.

⁷ Précisant dans l'en-tête la version PDF, la version Acrobat, la date de sortie ainsi que la correspondance, lorsqu'elle existe, avec le registre des formats PRONOM : <http://www.nationalarchives.gov.uk/PRONOM/Default.aspx>