

PDF/A: Revue des produits convertisseurs PDF/A aptes au mode batch

Quels sont les convertisseurs PDF/A courants? Combien coûtent-ils? Existe-t-il aussi des convertisseurs pour Linux? Sont-ils réellement aptes au mode batch? La présente étude répond à ces questions.

1. Introduction

Cette compilation de différents programmes pour la conversion de fichiers texte et image en PDF/A (convertisseurs) doit servir d'aide aux archives à deux égards: Premièrement, elle indique, selon les connaissances actuelles, quels convertisseurs aptes au mode batch existent pour la conversion de fichiers en PDF/A. Deuxièmement, elle peut servir de base pour une première évaluation par les différentes archives.

2. Caractéristiques

Parmi la multitude des convertisseurs existants, cinq ont été plus précisément analysés, et les caractéristiques suivantes en ont été retenues:

- Open source
- Licence libre / coûts de licence
- Aptitude au mode batch *i*
- Systèmes d'exploitation supportés

Toute l'analyse a été complétée par la conversion suivie d'une validation de différents documents de test:

- Photos (JPG & TIFF)
- Graphiques (PDF & PNG)
- Document Word *ii* (DOC)

3. Revue des produits

Lors du choix des programmes proposés, il a également été tenu compte de leur diffusion et d'un éventuel support. On a tenté en outre de proposer un choix aussi diversifié que possible.

Parce que les tests prennent beaucoup de temps, le CECO renonce à tester chaque nouvelle version séparément. En outre, il n'est pas possible pour nous d'identifier tous les convertisseurs existants. Sur demande, le CECO est toujours prêt à faire des mises à jour ou des ajouts.

3. Revue des produits

dmstools Rendering Service, de dmstools

(www.dmstools.de)

- Logiciel payant (EUR 4'400.- *i* sans maintenance ni support, pas de limite de volume).
- Ne convertit pas uniquement en PDF/A, mais également en PDF et TIFF.
- Afin d'éviter les coûts d'installation liés au test, les documents de test ont été créés d'une part à l'aide du système de test MailManager de dmstools, publiquement disponible et qui utilise ce service également pour la création de fichiers PDF/A. Les documents de test ont été convertis d'autre part directement chez dmstools, par Rendering Service. La qualité est très bonne, et le volume ne représente plus que 28% du volume original.
- La validation de tous les documents de test a réussi avec les trois programmes. A ce propos, le choix des documents de test (créés par dmstools Rendering Service via MailManager ou directement par dmstools Rendering Service) n'a joué aucun rôle.
- Le mode batch est assuré, selon le fournisseur, d'une part par un service basé sur un serveur. Lors de notre test, le Rendering Service a été testé par l'intermédiaire du système MailManager de dmstools et cela a parfaitement fonctionné. Le Rendering Service de dmstools offre d'autre part d'autres interfaces, par exemple API (seulement Windows), ligne de commande (Windows / Unix / Mac) et XML (indépendant de la plateforme).
- Le Rendering Service de dmstools ne fonctionne que sur Microsoft Windows Server 2003 et 2008, sur lesquels Office ou tout autre programme permettant d'afficher et d'imprimer des documents Office doit être installé. Ces diverses interfaces permettent cependant d'intégrer toutes les plateformes (p. ex. Unix / Mac).
- Ce programme peut décompresser tous les documents d'une archive (p. ex. fichiers ZIP et MSG) et les convertir en PDF/A.
- Le Rendering Service de dmstools peut être étendu par modules à des formats propriétaires et à des formats publics.
- Plusieurs serveurs peuvent être mis en parallèle afin d'assurer la tolérance aux pannes et pour répartir la charge.

A la fin de la revue du produit, le fournisseur a été contacté par courrier afin qu'il puisse prendre position face aux résultats présentés.

Position de dmstools concernant le dmstools Rendering Service

Réponse reçue: la rubrique concernant le Rendering Service de dmstools est correcte et dmstools n'a aucune correction à formuler. En complément, le contact suivant nous a encore été communiqué par mail:

Vous pouvez obtenir de plus amples informations concernant d'autres produits pour la création de PDF/A à la place de travail ou sur un serveur auprès de:	
dmstools AG Hindenburgstraße 73 74321 Bietigheim-Bissingen Allemagne WWW: http://www.dmstools.de	Personne de contact: Markus Wössner Tél. +49 (7142) 971334-0 Fax: +49 (7142) 971334-99 E-mail: info@dmstools.de

3. Produktreview

PDF/A Live!, de intarsys

(www.intarsys.de)

- Logiciel payant (€ 589.05 *i*, pas de limite de volume mais avec une limite de 25 conversions respectivement validations par lot).
- Convertit uniquement des fichiers PDF et des formats d'image (p.ex. JPG, TIF et PNG) en PDF/A. La qualité de la conversion JPG et PNG est très bonne, et le volume ne représente plus que 32% du volume original. La qualité de la conversion du PDF est très bonne, cependant, la taille du fichier augmente considérablement (153% de la taille originale). Dans la conversion du TIFF une erreur c'est produit lors du chargement du fichier [error loading image (I/O error reading image metadata!)]. Malheureusement, aucun message d'erreur est apparu et la conversion n'a pas été arrêtée. A la place du document, le message d'erreur a été converti. Avec le TIFF original (10,8 Mo) il n'y avait pas d'erreurs.
- La validation des documents de test a réussi. Lors de la validation avec PDF/A Manager, une seule erreur est apparue pour le documents de test PDF qui ont été convertis en PDF/A, indiquant que la date de création n'était pas synchronisée avec XMP.
- Le traitement par lot a été testé avec succès. L'aptitude au mode batch est assurée, selon le fournisseur, par la ligne de commande et par l'API. Autrement il est possible de faire tourner un service basé sur un serveur qui convertit les document d'un répertoire d'entrée.
- Fonctionne avec tous les systèmes d'exploitation usuels: Microsoft Windows (2000, XP, Vista), Linux et Mac OSX.
- Contient une validation.

A la fin de la revue du produit, le fournisseur a été contacté par courrier afin qu'il puisse prendre position face aux résultats présentés.

Position de intarsys concernant PDF/A Live! (en allemand)

Wir haben folgende Erläuterungen zu unseren Tests beziehungsweise Testdokumente und deren Behandlung durch PDF/A Live! von intarsys erhalten:

TIFF-Dokument: Unsere Analysen haben ergeben, dass im (komprimierten) TIFF-Dokument enthaltenen Metadaten (wie z.B. Erzeuger, Kamera, etc) an ungültigen Positionen referenziert werden. Es scheint offensichtlich so zu sein, dass beim Komprimieren des TIFFs die Byte-Position der Metadaten vom Kompressionsprogramm nicht an das verkleinerte Dokument angepasst wurden, sondern immer noch an die Position im Original-TIFF zeigen. Dieses ist ein Defekt im TIFF, der aber offensichtlich von "toleranteren" Konvertierungsprogrammen ignoriert wird. Unsere TIFF-Bibliothek wirft korrekterweise einen Fehler. Die fehlerhafte Position der Metadaten betrifft nicht die Anzeige des TIFFs, so dass das Vorgehen der anderen Konvertierer nicht generell falsch ist. Die Metadaten des original unkomprimierten TIFFs sind aber teilweise verloren.

PDF-Dokument: Das verwendete Dokument ist ein für die Druckvorstufe optimiertes PDF-File, bzw es basiert auf einer Druckertest-Datei. Diese verwendet den CMYK Farbraum in unkalibrierter Form (im Gegensatz zu RGB, der in digitalen Dokumenten weitaus häufiger benutzt wird). Wir verwenden für die PDF/A-konforme Darstellung unkalibrierter Farbräume frei verfügbare Farbraum-Definitionen. Dieser ist im Fall des CMYK-Farbraums 1,3 MB gross. Da dieser eingebunden werden muss, wächst das Dokument um genau diese Größe. Andere Konvertierer haben eigene bzw. lizenzierte Farbräume, die deutlich kleiner sind und manipulieren zusätzlich die Kompression des eingebundenen Bildes. Dies machen wir nicht im Falle eingebundener Bilder mit erlaubten Kompressionsverfahren.

i Le prix indiqué est valable pour la version Windows avec 'client batch'. Ceci permet la conversion et la validation avec max. 25 conversions / validations par lot (Une conversion / validation se réfère à un document ou bien par 10 pages, si le document a plus de 10 pages). Il existe également des licences de serveur, qui convertissent la pile complète, mais qui ont une limite de volume.

3. Revue des produits

LuraDocument PDF Compressor Enterprise avec Born Digital Module de LuraTech

(www.luratech.com)

- Logiciel payant (EUR 3'900.- *i* sans maintenance ni support, limite de volume *ii* fixée à 50'000 pages par an).
- Les documents de test ont été convertis en PDF/A par LuraTech à l'aide du compresseur PDF assisté du module Born Digital, la version de test n'étant pas conforme PDF/A à cause de son tampon transparent. Tous les documents de test ont été convertis avec succès par le module Born Digital. La qualité des résultats lorsque la compression des données est désactivée est très bonne, et le volume ne représente plus que 46% du volume original.
- La validation de tous les documents de test a réussi avec les trois programmes.
- L'aptitude au batch est assurée, selon le fournisseur, par le chargement et l'exécution automatique de «l'Automatic Job List Processing», à l'aide d'un fichier DAT. Cette possibilité n'a cependant pas été testée.
- Ne fonctionne que sur Microsoft Windows Server 2003 ou Server 2008 avec Office installé.

A la fin de la revue du produit, le fournisseur a été contacté par courrier afin qu'il puisse prendre position face aux résultats présentés.

Position de LuraTech concernant PDF Compressor Enterprise avec Born Digital Module

Réponse reçue: la rubrique concernant LuraDocument PDF Compressor est correcte, et LuraTech n'a aucune correction ni complément à formuler.

i Ces coûts concernent le modèle de licence limité à 50'000 pages par an de la version Advanced Enterprise avec le module Born Digital.
ii Des volumes de pages supplémentaires doivent être achetés si le besoin est supérieur à 50'000 pages par an..

3. Revue des produits

PDFCreator 0.9.9, de pdfforge *i*

(www.pdfforge.org)

- Programme Open Source gratuit, que les entreprises sont également autorisées à utiliser.
- Ne convertit pas uniquement en PDF/A, mais également en PDF, TIFF, PNG, JPG, BMP, PCX, PS et EPS.
- Tous les fichiers ont été convertis. La qualité des résultats est très bonne et le volume ne représente plus que 21% du volume original. Le «Windows Picture & Fax Viewer» est utilisé pour la conversion des images: dans ce programme, les réglages d'impression doivent être saisis manuellement. En conséquence, ces réglages ne conviennent pas au mode batch. Lorsque le Windows Picture & Fax Viewer est désactivé, le PDF Creator ne convertit plus les images.
- La validation des documents de test n'a réussi qu'avec Adobe Reader 9. Aucun document n'a passé la validation avec PDF/A Live! ni avec PDF/A Manager. PDF/A Live! a produit entre 2 et 10 erreurs.
- L'aptitude au traitement par lots par ligne de commande a été testée avec succès à l'aide d'un script batch sur les graphiques de test PDF. L'aptitude au mode batch n'est pas avérée pour les images, voir remarque ci-dessus.
- Ne fonctionne que sous Microsoft Windows (98, ME, NT, 2000, XP, 2003, Vista, 7) avec Office installé.
- Existe dans pratiquement toutes les langues.

A la fin de la revue du produit, le fournisseur a été contacté par courrier afin qu'il puisse prendre position face aux résultats présentés.

Position de pdfforge concernant PDFCreator 0.9.9

Aucune réponse reçue.

3. Revue des produits

3-Heights Document Converter Service, de PDF Tools

(www.pdf-tools.com)

- Logiciel payant (CHF 8'250.- *i* sans maintenance ni support, limite de volume *ii* fixée à 360'000 pages par an).
- Ne convertit pas uniquement en PDF/A, mais également en PDF et TIFF.
- Tous les fichiers ont été convertis. La qualité des résultats est très bonne et le volume ne représente plus que 43% du volume original.
- La validation des documents de test a réussi. Une seule erreur est apparue pour JPG, TIFF et PNG, lors de la validation avec PDF/A Manager, indiquant que la date de création n'était pas synchronisée avec XMP.
- L'aptitude au mode batch est assurée, selon le fournisseur, par Shell Tool d'une part (script batch par ligne de commande). La possibilité existe d'autre part de faire tourner le processus par API (Application Programming Interface) ou par un service basé sur un serveur.
- Le programme est une solution client serveur, étant entendu que le serveur doit être Microsoft Windows Server 2003 ou 2008 avec Office installé. Lorsque le service est basé sur un serveur, le client est indépendant du système d'exploitation.
- Le programme peut décompresser tous les documents d'une archive (p. ex. fichiers ZIP et MSG) et les convertir en PDF/A.
- Le 3-Heights Document Converter Service peut être adapté à des formats propriétaires à l'aide de modules programmables.
- Contient une validation.

A la fin de la revue du produit, le fournisseur a été contacté par courrier afin qu'il puisse prendre position face aux résultats présentés.

Position de PDF Tools concernant 3-Heights Document Converter Service

La prise de position reçue est jointe ci-dessous (en allemand). Elle contient en particulier les spécifications du contenu de la version du Document Converter Service pour poste de travail individuel (Small) et pour PME (Medium), qui sera disponible à fin juillet 2010.

i La licence permet à plusieurs archives de soumettre des documents à un serveur central pour conversion.

ii Des volumes de pages supplémentaires doivent être achetés si le besoin est supérieur à 360'000 pages par an.

3. Revue des produits

PDF/A Manager, de PDFTron

(www.pdftron.com)

- Logiciel payant (CLI: \$ 699.- ou SDK: \$ 799.- *i* sans maintenance ni support, pas de limite de volume).
- Convertit uniquement des fichiers PDF en PDF/A. En outre, le volume du fichier ne change que très peu (83% du volume original).
- La validation des documents de test a réussi avec tous les trois programmes.
- L'aptitude au mode batch par ligne de commande a été testée avec succès sur les graphiques de test PDF en utilisant un script batch. L'aptitude au mode batch de la version SDK est assurée, selon le fournisseur, par l'intermédiaire d'API.
- Fonctionne avec tous les systèmes d'exploitation usuels: Microsoft Windows (98, NT, 2000, XP, 2003, Vista), Linux et Mac OSX.
- Contient une validation.

A la fin de la revue du produit, le fournisseur a été contacté par courrier afin qu'il puisse prendre position face aux résultats présentés.

Position de PDFTron concernant PDF/A Manager

Réponse reçue: depuis juin 2010, PDFNet de PDFTron peut être acheté avec l'extension PDF/A. Ceci rend désormais la conversion en PDF/A possible. L'avantage par rapport au PDF/A Manager est la conversion de tous les types de documents usuels grâce à l'extension Convert. Le communiqué de presse correspondant (en anglais) est joint ci-dessous. Les indications de prix/performances détaillées peuvent être consultées sur www.pdftron.com/pdfnet.

4. Aperçu des produits

	Open source	Coûts de licence ⁱ	Documents de test ⁱⁱ						Σ ^{iv}	Aptitude au mode batch	Validation ⁱⁱⁱ	Systèmes d'exploitation
			.jpg 1.1 MB	.tif 3.5 MB	.pdf 2.9 MB	.png 3.7 MB	.doc 0.13 MB					
dmstools Rendering Service www.dmstools.de			✓ 1'086 KB	✓ 185 KB	✓ 234 KB	✓ 1'346 KB	✓ 300 KB	3.2 MB (28%)		✓✓✓		
PDF/A Live! www.intarsys.de			✓ 529 KB	? Fehler beim Laden	✓ 4'429 KB	✓ 563 KB	✗ nicht unterstützt	5.5 MB (72%)		✓✓✓		
LuraDocument PDF Compressor Enterprise mit Born Digital Modul www.luratech.com			✓ 1'025 KB	✓ 547 KB	✓ 2'399 KB	✓ 753 KB	✓ 492 KB	5.2 MB (46%)	Automatic Job List Processing	✓✓✓		
PDFCreator 0.9.9 www.pdfforge.org			✓ 807 KB	✓ 170 KB	✓ 216 KB	✓ 943 KB	✓ 219 KB	2.4 MB (21%)		✓✗✗		
3-Heights™ Document Converter Service www.pdf-tools.com			✓ 1'068 KB	✓ 187 KB	✓ 2'404 KB	✓ 902 KB	✓ 387 KB	4.9 MB (43%)		✓✓✓		
PDF/A Manager www.pdftron.com			✗ pas supporté	✗ pas supporté	✓ 2'408 KB	✗ pas supporté	✗ pas supporté	2.4 MB (83%)	/	✓✓✓		

konvert2pdfa_Tableaux_2010.11.doc

Etat: 02.12.2010

ⁱ En ce qui concerne les coûts de licence, c'est chaque fois l'équivalence quantité/prix la plus basse qui a été introduite dans l'évaluation. Il s'agit de 50'000 pages par an pour LuraTech et de 360'000 pages par an pour PDF Tools. Les autres fournisseurs ne font état d'aucune restriction équivalente.

ⁱⁱ Voici les résultats en termes de taille du fichier et la qualité des documents de test convertis. Toutefois, il ne contient aucune déclaration au sujet du traitement en mode batch et de leur conversion correcte vers PDF/A.

ⁱⁱⁱ Les documents PDF/A ont été validés par Adobe Reader 9, PDF/A Live! d'Intarsys Consulting GmbH et PDF/A Manager de PDFTRON. Dans ce contexte, c'est Adobe Reader 9 qui pose le moins d'exigences: La validation est considérée comme réussie si le document est reconnu comme PDF/A. PDF/A Live! fait la différence entre erreur et avertissement. Pour notre test, seules les erreurs étaient déterminantes. Le PDF/A Manager pose les plus hautes exigences à un PDF/A. Il ne connaît (globalement) que Fail ou Pass; cela malgré que la spécification PDF/A laisse une marge d'interprétation. Lorsque les documents de test n'ont pas tous passé la validation, une coche d'appréciation plus petite a été insérée.

^{iv} Les volumes de données obtenus sont à prendre comme exemples, car la qualité peut être paramétrée dans la plupart des programmes et le volume des fichiers varie en conséquence.

konvert2pdfa_Tableaux_2010.11.doc

Rc, 02.12.2010

Az

v0.1

Seite 1/1

3-Heights™ Document Converter Service

Der 3-Heights™ Document Converter Service ist die umfassendste Lösung der Produktpalette der PDF Tools AG. Sie zeichnet sich durch sehr weit gehende Funktionalitäten aus und hat im Grundpreis einen Durchsatz von 360'000 Seiten pro Jahr integriert. Deshalb kommt diese Lösung vor allem bei grösseren Unternehmen und Organisationen zum Einsatz.

Beispiele solcher Funktionalitäten sind:

- Konversion von Emails inkl. Anhänge
- Konversion von Webseiten
- Schrifterkennung mittels OCR
- Qualifizierte Digitale Signaturen
- Skalierbarkeit auf Multiprozessorsysteme und Servercluster

PDF Tools AG hat erkannt, dass diese Lösung für KMUs zu umfangreich sein kann und wird deshalb zwei Varianten der Maximallösung (Large) des Document Converter Services anbieten.

Nachfolgend die inhaltlichen Spezifikationen zu der Einzelplatz- (Small) und der KMU-Version (Medium) des Document Converter Service, welche bis Ende Juli 2010 bereitgestellt werden.

a) Lösung 'Small' für den Desktop:

Der Benutzer soll die folgenden Funktionen interaktiv benutzen können:

- Dokumente (PDF, TIFF, JPEG usw.) anzeigen
- Dokumente zusammenfügen, Reihenfolge und Orientierung der Seiten ändern, Seiten löschen
- PDF/ A direkt aus Applikationen und nach der Bearbeitung erzeugen
- PDF Dateien auf PDF/A Konformität prüfen
- Ein ganzes Verzeichnis mit Dokumenten auf einmal konvertieren (Batchverarbeitung)
- Qualifizierte Digitale Signaturen (Einzelsignaturen)
- Optional: OCR

Preis: 850.00 CHF (exkl. OCR)

b) Lösung 'Medium' für KMU's:

- Der Benutzer soll die folgenden Funktionen benutzen können:
- Funktionen aus 'Small'
- Zusätzlich einfacher Windows Service Prozess mit überwachten Verzeichnissen mit allen Konversionsfunktionen (Office to PDF/A, PDF to PDF/A, Image to PDF/A usw.)
- Qualifizierte Digitale Signaturen
- Optional: OCR

Preis: 3'500.00 CHF (exkl. OCR)

Damit wird PDF Tools AG ab August 2010 für jedes Budget ein Angebot mit optimalem Preis/Leistungsverhältnis anbieten können.

PDF Tools AG

Geerenstrasse 33

8185 Winkel

Schweiz

Tel.: +41 43 411 44 51

Fax: +41 43 411 44 55

Email: pdfsales@pdf-tools.com

PDFTron Launches a New Version of PDFNet SDK featuring Universal Document Conversion

Vancouver, Canada, June 1st, 2010

PDFTron Systems Incorporated, a global provider of industry-leading, high-performance electronic document format technology, today unveiled PDFNet SDK 5, the latest new major release of PDFTron's flagship PDF development component for advanced, high-quality PDF creation, content extraction, editing, viewing and printing. This release features many important enhancements and launches two brand new modules adding support for universal document conversion and more.

PDFNet SDK is an advanced PDF toolkit used by developers requiring a higher level of PDF integration, performance and control within their commercial applications and enterprise workflows. With PDFNet it is easy to add capabilities such as PDF viewing, PDF to image conversion, PDF content extraction and advanced PDF editing operations.

With PDFNet SDK 5, the toolkit's already powerful set of features is expanded to include universal document conversion and enhanced PDF/A conversion as part of two brand new add-on modules to PDFNet: Convert Add-on and PDF/A Add-on. The new modules provide developers the option to leverage the same toolkit for all types of PDF processing capabilities as well as for fast and reliable universal document conversion.

The Convert Add-on offers conversion from any printable document format to PDF, XPS, EMF, or SVG using a very simple universal conversion API based on the XPS print path, enabling superior conversion output on Windows 7 and Vista, compared to any GDI/PostScript driver. To provide the highest quality output PDFNet implements a number of direct converters that preserve text, vector graphics, hyperlinks, colors and fonts. PDFNet can also seamlessly integrate with Microsoft Office to maintain metadata, text, links when converting from Word, Excel, Outlook, or PowerPoint to PDF or XPS. The universal conversion API is very easy to use and applications using this functionality will not need to separately distribute, install, or register any additional files.

The PDF/A Add-on enables direct PDF to PDF/A conversion, supporting both PDF/A-1a and PDF/A-1b. Included in this add-on are various customization options and features to check compliance and produce reports. Because PDF/A Add-on seamlessly integrates with the PDF creation, viewing, printing, and document manipulation features of PDFNet, organizations can quickly implement their workflows without having to deal with a mix of incompatible components.

"The extension of PDFNet functionality through the new modules featuring universal document conversion allows our customers to get even more accomplished with just one toolkit, while not bloating the component or price for those that do not need the all-inclusive functionality" says Catherine Andersz, Director of Business Development. "The new modules provide all the flexibility and performance without any unnecessary complexity."

Using the new version of PDFNet SDK, developers will also benefit from many other new major enhancements, including a new fast and efficient Print API as well as a greatly improved PDF Viewing Control, now featuring a set of built-in tools allowing developers to navigate, markup, annotate, and fill-in PDF forms out-of-the box.

PDFNet product information, sample code, and full-featured trial versions are available for immediate download at <http://www.pdftron.com/pdfnet/index.html>.

PDFNet SDK is available as a .NET component as well as a cross-platform Java and C/C++ library for Windows, Linux, Mac OSX and other platforms.

As always, PDFNet SDK supports all versions of the PDF Specification, including the PDF ISO32000 as well as Adobe® Acrobat® 9.x documents.

About PDFTron Systems Inc.

PDFTron Systems Inc. (<http://www.pdftron.com>) is a global provider of industry-leading, high-performance document processing and conversion technology related to PDF, XPS, SVG and other graphic file formats. Our flagship product, [PDFNet SDK](#), provides organizations a cost-effective, reliable and easy way to seamlessly integrate all types of PDF capabilities and document conversion functionality with complex enterprise-wide document management and publishing workflows, interactive client applications as well as high-traffic server-based applications. Forward-thinking, customer-centric, and quality-obsessed are some of the traits that have turned PDFTron into the primary resource for major technology, engineering, and healthcare companies as well as financial, educational and government institutions world-wide.

Contact

PDFTron Press
Tel: 604-730-8989
Fax: 604-676-2477
Email: press@pdftron.com