

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

Quels sont les convertisseurs PDF/A courants ? Combien coûtent-ils ? Que peut-on en attendre ? Pour quel usage sont-ils appropriés ? La présente étude répond à ces questions.

Table des matières

1. Introduction.....	1
2. Convertisseurs testés.....	1
3. Corpus de test.....	2
4. Analyse des convertisseurs.....	2
5. Description détaillée des convertisseurs.....	3
6. Aperçu des résultats de conversion et caractéristiques.....	10

1. Introduction

Cette compilation de différents programmes de conversion de fichiers PDF, Office et image en fichiers 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 ces trois catégories de formats sources ; deuxièmement, elle peut servir de base à une première évaluation par les différentes archives et autres services versants. Cette étude est la version entièrement remaniée et mise à jour de la première étude sur les convertisseurs PDF/A de 2010¹.

2. Convertisseurs testés

Les tests étant très coûteux, cette étude ne sera pas mise à jour régulièrement. Elle représente une image instantanée prise à la fin 2016. Le CECO ne peut pas montrer tous les convertisseurs existants. Les mises à jour ou compléments peuvent être annoncés au CECO.

L'étude est complétée d'une liste séparée² (en allemand uniquement) comprenant d'autres convertisseurs qui ne correspondent pas aux trois catégories de format d'entrée ou qui n'ont pas pu être considérés dans l'étude faute de temps. Cette liste sera actualisée en permanence.

Pour établir cette sélection de convertisseurs, les connaissances préalables du CECO issues de l'étude précédente, les indications en provenance des archives ainsi que la diffusion et la diversité des convertisseurs ont été prises en compte. Sept convertisseurs de sept fabricants différents ont été testés. Ils sont décrits en détail au chapitre 5.

- Adobe : Acrobat Pro DC
- Callas : pdfaPilot CLI 6
- dmstools: dmstools Rendering Service
- Foxit: PDF Compressor Enterprise avec module Born Digital
- Neevia: Document Converter Pro
- PDF Tools AG : 3-Heights Document Converter SME
- PDFTron: DocPub & PDF/A Manager

¹ <http://kost-ceco.ch/cms/download.php?31137eb7707926114dad7654f982a411>

² http://www.kost-ceco.ch/cms/index.php?pdfa_konverter_add_de

3. Corpus de test

Le corpus de test était constitué de divers formats de fichiers de différente qualité. La vitesse et la robustesse ont été évaluées au moyen de formats de fichiers rares ainsi que de fichiers invalides et endommagés. Le corpus de test comprenait 342 fichiers, 288 fichiers pouvant être décrits comme réalistes, c'est-à-dire courants et non endommagés :

- Documents PDF, y c. PDF/A : 59
- Documents MS Office :
 - Documents Word : 35
 - Présentations PowerPoint : 18
 - Fichiers Excel : 18
 - Mails : 20
- Documents OpenOffice : 14
- Autres documents texte (TXT, RTF, HTML) : 10
- Images TIFF : 47
- Images JPEG : 20
- Images JPEG2000 : 23
- Autres images (BMP, PNG, GIF) : 24

La qualité variable et la taille du corpus de test ne permettent de faire aucune affirmation fiable quant à la performance. Pour cela, il faudrait beaucoup plus de fichiers pouvant être en principe convertis par tous les convertisseurs.

4. Analyse des convertisseurs

Parmi la multitude de convertisseurs existants, sept ont été analysés en détail. L'ensemble du corpus de test a été converti avec chaque convertisseur en PDF/A-2b. Les résultats ont ensuite été analysés en fonction des critères suivants :

- Validité : quel pourcentage de tous les fichiers réalistes³ a réussi la double validation⁴ avec KOST-Val⁵ (à l'exclusion de la validation image) ?
- Qualité : artéfacts de compression visibles, le résultat contient-il toutes les informations ?
- Diversité : quel pourcentage de tous les fichiers réalistes³ a été converti ?
- Vitesse et robustesse⁶ : répartition dans les cinq catégories suivantes au moyen des impressions obtenues lors du test
 - 100 % = les deux bons (moins de 30 minutes et aucun problème de robustesse)
 - 90 % = en moyenne suffisant à bon
 - 80 % = en moyenne suffisant
 - 70 % = en moyenne mauvais à suffisant
 - 60 % = les deux mauvais (bien au-delà d'une heure et plusieurs blocages)

De plus, nous avons consigné les caractéristiques suivantes :

- Coûts et modèle de licence
- Systèmes d'exploitation supportés et aptitude au mode batch⁷

Les coûts ne comprennent que les frais d'acquisition proprement dits, à l'exclusion des coûts de maintenance et de support. Ces derniers se montent en général à 20 % des coûts d'acquisition. La conversion étant une tâche à long terme, le CECO recommande en principe cet investissement supplémentaire afin de pouvoir toujours bénéficier des versions les plus actuelles.

³ Voir définition ci-dessus : fichiers courants non endommagés.

⁴ Dans la double validation, il suffit qu'un des deux validateurs désigne le résultat comme valide. Étant donné que KOST-Val utilise les validateurs de PDF Tools AG et PDFtron, lors de la validation de ces convertisseurs, la valeur moyenne entre la double validation et le validateur concurrent a été calculée.

⁵ Au moment d'effectuer ces tests, veraPDF se trouvait encore à l'état de prototype et en phase de test. Le CECO intégrera veraPDF dans KOST-Val dès que nos exigences en matière de qualité seront satisfaites et que veraPDF aura terminé la phase test.

⁶ La rapidité et la robustesse ont été volontairement évaluées ensemble étant donné que les deux critères s'influencent mutuellement. Une séparation n'est donc pas judicieuse.

⁷ Le traitement en mode batch ou par lot désigne le traitement séquentiel, non interactif des tâches.

Aucun critère propre n'a été défini concernant la convivialité étant donné que celle-là dépend fortement de l'aptitude au mode batch et de l'implémentation⁸.

Nous avons testé les convertisseurs dans leur configuration standard. Il est bien entendu possible d'adapter chaque convertisseur professionnel aux besoins particuliers. Le CECO a intentionnellement testé la configuration standard étant donné que c'est celle que le fabricant a manifestement jugée optimale.

Le CECO estime la qualité plus importante que la taille du fichier. Nous ne faisons pas de commentaire sur la taille du fichier vu que la valeur optimale est difficile à définir. Une taille de fichier très réduite peut indiquer une compression extrême, une très grande taille peut signifier que le texte est intégré en tant qu'image, les deux situations n'étant pas idéales.

En plus de l'étude, les archives membres du CECO recevront, sur demande, un aperçu détaillé de l'analyse ainsi que des réponses à des questions ciblées.

5. Description détaillée des convertisseurs

Chaque convertisseur est décrit en détail ci-après. Après avoir passé en revue les produits, nous nous sommes adressés aux fabricants afin qu'ils puissent donner leur avis⁹ sur les résultats obtenus lors des tests.

a) Adobe : Acrobat Pro DC

(<https://acrobat.adobe.com/ch/fr/acrobat/acrobat-pro.html>)

- Le logiciel coûte 656,65 CHF, sans maintenance ni support.
- Il n'y a pas de limite de volume et donc pas de coûts supplémentaires.
- Acrobat Pro DC tourne sur un serveur Windows, client Windows ainsi que sur Mac OS X
- Les fichiers à convertir peuvent être transmis via une action à établir.
- 87 % des fichiers convertis ont réussi la double validation PDF/A avec KOST-Val. La validité se situe donc dans la fourchette supérieure. Ce sont principalement des fichiers PDF et PDF/A qui n'ont pas réussi la validation avec KOST-Val étant donné que le validateur intégré n'a pas identifié les parties problématiques comme étant des erreurs et qu'il ne les a donc pas corrigées.
- Le taux de satisfaction des exigences de qualité de 95 % est conforme aux prévisions.
- La diversité des documents convertis de 78 % est dans la moyenne.
- La vitesse et la robustesse sont jugées mauvaises. Le traitement des fichiers PDF prend particulièrement beaucoup de temps.

Recommandation : nous pouvons recommander Acrobat Pro DC d'Adobe si l'on doit convertir manuellement des documents en PDF/A.

Position d'Adobe concernant l'évaluation d'Acrobat Pro DC :

Nous pouvons accepter les résultats des tests.

Nous souhaiterions qu'à l'avenir on utilise des validateurs neutres du point de vue des convertisseurs afin de pouvoir mieux comparer.

⁸ Un outil purement en ligne de commande est peu commode par exemple pour un utilisateur non averti, cependant il ne représente aucun problème s'il s'agit juste d'exécuter un script.

⁹ Les avis des fabricants ont été formulés en allemand et traduits par le CECO. Les avis originaux se trouvent dans la version allemande : <http://kost-ceco.ch/cms/download.php?8e7cdab5b9e953bb2508c1757e5a5c88>

b) Callas : pdfaPilot

(www.callassoftware.com/fr/produits/pdfapilot)

- pdfaPilot CLI coûte 4 499 EUR, sans maintenance ni support.
- Il n'y a pas de limite de volume et donc pas de coûts supplémentaires.
- pdfaPilot CLI est une application en ligne de commande qui est disponible sous Windows (32 et 64 bits), Linux (32 et 64 bits), Mac OS X, IBM AIX et Sun Solaris.
- pdfaPilot existe en trois versions aptes au mode batch. En plus de pdfaPilot CLI (en ligne de commande) qui est décrit ici et que nous avons testé, il existe pdfaPilot Server et pdfaPilot SDK (interface de programmation).
- 91 % des fichiers convertis ont réussi la double validation PDF/A avec KOST-Val. La validité est donc bonne.
- Le taux de satisfaction des exigences de qualité de 98 % est conforme aux prévisions
- La diversité des documents convertis de 78 % est dans la moyenne. Le pourcentage est particulièrement bas pour la conversion des images en PDF/A.
- La vitesse et la robustesse sont jugées suffisantes. Les fichiers du CLI ont dû être transmis un à un, faute de quoi une grande partie n'est pas convertie en raison des cas problématiques.

Recommandation : nous pouvons recommander pdfaPilot de Callas lorsqu'il s'agit surtout de conversions Office et PDF.

Position de Callas concernant l'évaluation de pdfaPilot :

1. Versions aptes au mode batch

En plus des solutions mentionnées, il existe également pdfaPilot Desktop, une application pour Windows ou Macintosh avec laquelle on peut traiter des fichiers par lot dans un dossier. Les résultats sont rangés à côté des originaux dans un sous-dossier.

2. Résultats de conversion

Nous estimons que le 9 % d'écart calculé lors de la validation est encore plutôt élevé pour respecter la norme PDF/A. Nous travaillons donc, conjointement avec d'autres fournisseurs de validateurs PDF/A et sous le patronage de la PDF Association, à des améliorations dans le cadre du projet veraPDF. D'une part, une TechNote corrigera les différences d'interprétation de la norme PDF/A encore présentes, d'autre part un validateur open source fournira une base fiable pour l'évaluation des résultats de conversion. Ce validateur open source ainsi que les fichiers de test également mis au point dans le corpus de test de veraPDF constituent aujourd'hui déjà une partie essentielle de notre propre assurance qualité qui assure une compatibilité complète avec veraPDF.

c) dmstools : dmstools Rendering Service

(www.dmstools.de/rendering-service.html)

dmstools Rendering Service existe sous deux formes différentes : la version serveur et la nouvelle version Workstation parue en mars 2017. Les tests de fin 2016 ont été entièrement réalisés avec la version serveur. Étant donné que les membres du CECO sont plus intéressés par la version Workstation, la présente étude l'a également prise en compte. Elle n'a pu être testée que brièvement sous la forme d'un prototype. Les résultats de conversion sont identiques avec la version serveur vu que le même Rendering Service est utilisé.

- Le logiciel pour serveur coûte 4 400 EUR, sans maintenance ni support.
 - La version Workstation coûte 2 400 EUR, sans maintenance ni support.
- Il n'y a pas de limite de volume et donc pas de coûts supplémentaires.
- dmstools Rendering Service tourne sur un serveur Windows. Le convertisseur peut être utilisé depuis n'importe quelle plateforme de système d'exploitation via XML ou un service web. La version serveur supporte deux conversions simultanées par serveur.
 - dmstools Rendering Service for Workstation tourne actuellement sur Windows 10, mais Windows 8 et Windows 7 sont en préparation. La version Workstation supporte une conversion à la fois.
- Les fichiers à convertir peuvent être transmis via une multitude d'interfaces, notamment d'une API (interface de programmation), d'un service web (SOAP), en ligne de commande, d'un dossier actif et depuis XML.
 - Dans le cas de la version Workstation, toutes les interfaces sont à disposition à l'exception du service web. Pour le bref test, nous avons utilisé un dossier actif.
- 99 % des fichiers convertis ont réussi la double validation PDF/A avec KOST-Val. La validité est donc bonne.
- Le taux de satisfaction des exigences de qualité de 97 % est conforme aux prévisions.
- La diversité des documents convertis de 96 % est bonne.
- La vitesse et la robustesse sont jugées bonnes.

Recommandation : Nous pouvons recommander la version serveur de dmstools Rendering Service de dmstools à tous ceux qui ont la possibilité d'exploiter ce convertisseur sur un serveur Windows. Nous pouvons également recommander dmstools Rendering Service for Workstation de dmstools pour tous ceux qui préfèrent la solution poste de travail Windows.

Position de dmstools concernant l'évaluation de dmstools Rendering Service

Nous vous remercions d'avoir pris en compte notre produit dans votre test.

Le test a porté sur la version 2.17 avec la configuration standard qui représente un compromis entre qualité et taille de fichier.

Entre-temps est parue la version 2.18 du produit qui contient encore une fois de nombreuses améliorations dans les domaines testés.

d) Foxit: PDF Compressor Enterprise

(<https://www.foxitsoftware.com/fr/products/luratech-pdf-compressor/>)

- PDF Compressor Enterprise avec le module Born Digital coûte 1 300 EUR pour 100 000 pages par an, sans maintenance ni support.
- Il y a une limitation de volume de 100 000 pages par an. Son dépassement engendre des coûts supplémentaires au prorata.
- PDF Compressor Enterprise avec le module Born Digital tourne sur un serveur Windows ou sur un client Windows 64 bits.
- Les fichiers à convertir peuvent être définis et démarrés via une interface utilisateur (GUI).
- 97 % des fichiers convertis ont réussi la double validation PDF/A avec KOST-Val. La validité est donc bonne.
- Le taux de satisfaction des exigences de qualité de 89 % est inférieur aux prévisions. 9,1 % des fichiers présentent une qualité faible et 0,9 % une qualité insuffisante¹⁰. Nous déplorons en particulier l'utilisation de la technologie de compression MRC, y compris JBIG2, sur des images qui ont déjà été comprimées. Le CECO recommande de ne pas utiliser JBIG2 pour le moment¹¹.
- La diversité des documents convertis de 72 % est dans la moyenne.
- La vitesse et la robustesse sont jugées suffisantes à bonnes.

Recommandation : Nous recommandons PDF Compressor Enterprise avec le module Born Digital de Foxit principalement pour un service spécialisé dans la numérisation si le client autorise l'utilisation de JBIG2 et si la taille du fichier doit être extrêmement réduite. Autre ombre au tableau, la limitation du volume et les coûts supplémentaires qui en résultent le cas échéant.

Position de Foxit concernant l'évaluation de PDF Compressor Enterprise y c. le module Born Digital :

Les faits sont correctement présentés, par exemple que le compresseur PDF tourne sur un système Windows 64 bits via une interface utilisateur.

En ce qui concerne les prix, ma collègue vous a indiqué une version « typique » et souvent utilisée du compresseur PDF et ce prix est correct. Cependant, nous proposons des modèles de licence très flexibles dans lesquels nous essayons d'offrir à chaque client la version qui lui convient le mieux. Pour des volumes importants, les licences sont naturellement plus chères.

Nous ne pouvons suivre l'évaluation qualitative que partiellement parce que nous n'avons pas pu disposer de l'ensemble de la batterie de tests pour examiner d'éventuels problèmes dans des fichiers individuels.

Nous cautionnons JBIG2 en tant que norme ISO, mais même si des clients ne veulent pas l'utiliser, par exemple en raison de la directive allemande TR Resiscan, nous obtenons encore une compression très élevée avec TIFF pour la couche dite noir/blanc avec notre compression 3 couches.

Nous ne saisissons pas complètement votre recommandation. Les services spécialisés dans la numérisation représentent une grande partie de notre clientèle et le compresseur PDF constitue une part de marché importante dans l'espace germanophone. Il convient cependant également pour toutes les autres branches et il y a en particulier de nombreux utilisateurs satisfaits auprès des autorités ainsi que dans les banques et assurances.

¹⁰ La qualité est insuffisante lorsque des informations sont perdues. Cela est doublement déduit du degré de satisfaction des exigences.

¹¹ http://kost-ceco.ch/cms/index.php?jbig2-compression_de

e) Neevia: Document Converter Pro

(www.neevia.com/convert-to-pdf)

- Document Converter Pro coûte 399 \$, sans maintenance ni support.
- Il n'y a pas de limite de volume et donc pas de coûts supplémentaires.
- Document Converter Pro tourne sur un serveur Windows ou sur un client Windows 32 ou 64 bits.
- On configure et démarre Document Converter Pro principalement via une interface utilisateur (GUI). En cas d'installation sur un serveur, Document Converter Pro peut également être utilisé comme un service Windows.
- 90 % des fichiers convertis ont réussi la double validation PDF/A avec KOST-Val. La validité est donc bonne.
- Le taux de satisfaction des exigences de qualité de 100 % est conforme aux prévisions.
- La diversité des documents convertis de 78 % est dans la moyenne.
- La vitesse et la robustesse sont jugées mauvaises à suffisantes. La vitesse en particulier ralentit avec le nombre de documents convertis.

Recommandation : Nous recommandons Document Converter Pro de Neevia en particulier pour les plus petites exploitations qui ont besoin d'une interface utilisateur Windows.

Position de Neevia concernant l'évaluation de Document Converter Pro :

Neevia ne nous a pas transmis sa prise position.

f) PDF Tools AG : 3-Heights Document Converter SME

(<http://www.pdf-tools.com/pdf20/fr/produits/pdf-converter-validation/document-converter/>)

- L'édition SME coûte 2 750 CHF, sans maintenance ni support.
- 72 000 pages par an sont comprises dans le prix. Le dépassement de cette limite engendre des coûts supplémentaires au prorata.
- 3-Heights™ Document Converter SME (Small-Medium Enterprise) tourne sur un serveur Windows ou sur un client Windows 64 bits. En cas d'installation sur un serveur Windows, le convertisseur peut être utilisé depuis n'importe quelle plateforme de système d'exploitation.
- Les fichiers à convertir peuvent être transmis dans la version SME via des répertoires surveillés, en ligne de commande, via une API ainsi que via l'extension Shell pour Explorer de Windows, y compris via des interfaces graphiques (GUI). L'édition entreprise permettrait d'utiliser encore d'autres interfaces comme un service web.
- 91 % des fichiers convertis ont réussi la double validation PDF/A avec KOST-Val¹². La validité est donc bonne.
- Le taux de satisfaction des exigences de qualité de 98 % est conforme aux prévisions.
- La diversité des documents convertis de 87 % est bonne.
- La vitesse et la robustesse sont jugées bonnes.

Recommandation : Nous recommandons 3-Heights™ Document Converter SME de PDF Tools AG à tous, en particulier aux petites et moyennes exploitations. Les exploitations plus grandes peuvent envisager l'édition entreprise. Seule ombre au tableau, le modèle de licence et les coûts supplémentaires qui en résultent.

Position de PDF Tools AG concernant l'évaluation de 3-Heights™ Document Converter :

Nous remercions le CECO d'avoir examiné consciencieusement la qualité des fichiers générés. Nous pouvons comprendre les résultats dans les conditions de tests choisies. Toutefois, il se trouve que pour nos vérifications de qualité, nous utilisons un autre jeu de validateurs.

Il s'agit principalement d'Adobe Preflight et de veraPDF, dont les fabricants, conjointement avec PDF Tools AG, travaillent activement à l'harmonisation des validateurs au PDF/A Competence Center.

Avec ce jeu de validateurs, les résultats des tests de conformité sont par conséquent nettement meilleurs.

¹² Parce que PDF Tools AG produit également un des deux validateurs dont se sert KOST-Val pour la double validation, on utilise la valeur moyenne entre la double validation et la validation par l'autre validateur employé, PDFTron, pour corriger une possible altération du jugement.

g) PDFTron : DocPub & PDF/A Manager

(www.pdftron.com/docpub/index.html & www.pdftron.com/pdfamanager/index.html)

DocPub, qui convertit divers fichiers en PDF, a été testé conjointement avec PDF/A Manager, qui effectue la conversion du PDF en PDF/A. De plus, PDFTron offre avec PDFNet SDK une série de développements logiciels qui permettrait de convertir directement des fichiers en PDF/A.

- On n'a pas pu nous indiquer les coûts ni le modèle de licence pour DocPub et PDF/A-Manager étant donné qu'ils dépendent fortement du domaine d'utilisation et de la manière dont ils sont intégrés.
- DocPub et PDF/A Manager tournent sur Windows 7, 2008, Vista, XP, 2003, 2000, NT, 98 ainsi que sur Mac OSX et Linux.
- Les fichiers à convertir peuvent l'être en ligne de commande.
- 94 % des fichiers convertis ont réussi la double validation PDF/A avec KOST-Val¹³. La validité est donc bonne.
- Le taux de satisfaction des exigences de qualité de 97 % est conforme aux prévisions.
- La diversité des documents convertis de 83 % est bonne.
- La vitesse et la robustesse sont jugées bonnes.

Recommandation : Nous pouvons recommander la combinaison de DocPub et PDF/A Manager de PDFTron à tous ceux qui préfèrent une solution console.















Position de PDFTron concernant l'évaluation de DocPub & PDF/A Manager :




PDFTron ne nous a pas transmis sa prise de position.

¹³ Parce que PDF Tron produit également un des deux validateurs dont se sert KOST-Val pour la double validation, on utilise la valeur moyenne entre la double validation et la validation par l'autre validateur employé, PDF Tools AG, pour corriger une possible altération du jugement.

6. Aperçu des résultats de conversion et caractéristiques

Les résultats des conversions et les caractéristiques des convertisseurs sont résumés dans cet aperçu. La pondération est une proposition du CECO. Vous pouvez télécharger le fichier Excel¹⁴ afin d'effectuer votre propre pondération.

Étude 2016 du CECO sur les convertisseurs	Adobe Acrobat Pro DC	Callas pdfaPilot	dmstools dmstools Rendering Service	Foxit PDF Compressor Enterprise	Neevia Document Converter Pro	PDF Tools AG 3-Heights Document Converter SME	PDFTron DocPub & PDF/A Manager	Pondération
Coûts	CHF 656,65	EUR 4 499.-	(EUR 4 400.-) EUR 2 400.-	EUR 1 300.-	\$ 399.-	CHF 2 750.-	Ne peut être indiqué de manière forfaitaire	-
Modèle de licence	-	-	-	Nombre de pages limité	-	Coûts suppl. dès 72 000 p/y	Ne peut être indiqué de manière forfaitaire	-
Systèmes d'exploitation								-
Aptitude au mode batch								-
Validité	87 %	91 %	99 %	97 %	90 %	91 %	94 %	2
Qualité	95 %	98 %	97 %	89 %	100 %	98 300	97 %	3
Diversité	78 %	78 %	96 %	72 %	78 %	87 %	83 %	2
Vitesse & robustesse	60 %	80 %	100 %	90 %	70 %	100 %	100 %	1
Total pondéré	84 %	89 %	98 %	87 %	88 %	94 %	93 %	8

Légende de l'illustration Aptitude au mode batch :  En ligne de commande  Interface de programmation  Service  Interface utilisateur

¹⁴ <http://www.kost-ceco.ch/cms/download.php?5e0305433ad86c07a850ac15798d0c21>