

Validateurs PDF/A

Les membres du CECO le savent depuis longtemps : réclamer des formats adaptés pour l'archivage sans en vérifier la conformité ne représente que la moitié du travail.

C'est la raison pour laquelle le CECO ne se contente pas de mettre à jour continuellement son Catalogue des formats de données d'archivage [Cfa](#) , mais qu'il s'active depuis des années à la validation de formats. Les réflexions de base comme celles développées dans l'étude « [Reconnaissance et validation de format - Théorie et pratique](#) » de 2012 constituent le fondement théorique qui trouve sa concrétisation dans le validateur multiformat [KOST-Val](#) .

La validation du PDF/A, format essentiel pour l'archivage numérique, revêt une importance particulière. C'est pour cette raison que le CECO a testé en 2010 déjà sept validateurs et analysé les résultats pour les publier dans une étude. Après sept ans, il semble que le moment est venu de remanier entièrement cette analyse. Nous avons le plaisir de vous présenter la nouvelle étude du CECO sur les validateurs de PDF/A. Cette étude comprend encore cinq produits importants du marché parmi lesquels le nouveau validateur open source veraPDF développé dans le cadre du projet PREFORMA.

Une analyse complète de validateurs de PDF/A nécessite un corpus de test global, correct et si possible non publié de fichiers qui permettent une analyse séparée de toutes les exigences qui nous intéressent. L'ensemble de test Bavaria a mis un tel corpus de test à disposition en 2010. Cet outil est entre-temps devenu non seulement obsolète, mais il est en plus également trop bien connu. Développer un nouveau corpus de test dépasse les possibilités actuelles du CECO, raison pour laquelle la nouvelle version de l'étude sur les validateurs de PDF/A se base sur un banc d'essai de validateurs (benchmarking). Nous avons utilisé un corpus de 2980 fichiers PDF réels différents provenant de l'environnement du CECO et du groupe de travail nestor sur la reconnaissance des formats qui ont servi à tester tous les validateurs analysés.

Les aspects suivants ont été consignés (en plus de la version, du testeur et de la date du test) :

- Prix : prix du produit y compris contrat de maintenance pour un an.
- Vitesse : durée de la validation complète dans l'environnement de test.
- Robustesse : nombre de sorties incontrôlées au cours d'une validation complète.
- Consensus : dans 82.58% des cas, en juin 2017 tous les validateurs étaient d'accord et dans seulement 3.52% des cas il n'y avait pas de résultat net (2 contre 2). L'écart par rapport à la majorité des autres validateurs a été consigné dans le résultat du test qu'il soit valide ou invalide. Le résultat indique en plus la répartition entre documents valides et invalides.
- Précision : étant donné qu'il s'est avéré impossible de réaliser un mappage automatisé des messages d'erreur, nous avons effectué un contrôle manuel de la qualité sur plus de 30 fichiers du test qui n'avaient pas réussi la validation en juin 2017. Nous avons consigné le pourcentage des messages d'erreur concordants.

Les résultats de l'analyse sont rassemblés dans le tableau ci-après. Veuillez noter que nous avons testé trois versions différentes de veraPDF.

Validateurs PDF/A 2017	Callas: pdfaPilot	PDF Tools: 3Heights PDF Validator	PDFTron: PDF/A Manager	veraPDF (v1.6.1)	veraPDF (v1.8.4)	veraPDF (v1.10.4)	Intarsys PDF/A Live!
Prix Gratuit: Gratuit Bas: 1 - 499 CHF Modéré: 500 - 999 CHF Élevé: > 999 CHF	Élevé ³ EUR 5'399.-	Bas CHF422.-	Modéré ³ USD 699.-	Gratuit CHF 0.00	Gratuit CHF 0.00	Gratuit CHF 0.00	Modéré ³ de EUR 474.10 à EUR 1'200.-
Vitesse Très bien: ≤ 30 minutes Bien: 31 - 120 minutes Suffisant: 121 - 240 minutes Insuffisant: > 240 minutes	Bien 1:58:50	Très bien 0:18:00	Très bien 0:13:38	Insuffisant 8:30:20	Insuffisant 9:22:45	Suffisant 3:28:33	Très bien 0:15:59
Robustesse Très bien: ≤ 5 sortie incontrôlée Bien: 6 - 10 sortie incontrôlée Suffisant: 11-30 sortie incontrôlée Insuffisant: > 30 sortie incontrôlée	Très bien 0	Très bien 1	Très bien 3	Insuffisant 49	Insuffisant 55	Insuffisant 48	Bien 6
Consensus Très bien: < 1.0% Écart Bien: 1.0-4.9% Écart Suffisant: 5.0-9.9% Écart Insuffisant: ≥ 10.0% Écart	Bien	Très bien	Suffisant	Bien	Insuffisant	Suffisant	Bien
Écart total	2.38%	0.87%	5.74%	4.90%	19.53%	5.91%	1.51%
Reste valide	0.00%	0.81%	5.00%	2.28%	10.23%	3.22%	1.17%
Reste invalide	2.38%	0.07%	0.74%	2.62%	09.30%	2.68%	0.34%
Précision Très bien: ≥ 90% message d'erreur approprié Bien: 80-89% message d'erreur approprié Suffisant: 70-79% message d'erreur approprié Insuffisant: < 70% message d'erreur approprié	Bien 80.89%	Très bien 92.67%	Bien 82.11%	Bien 86.44%	Insuffisant 26.56%	Bien 83.78%	Bien 80.06%
Version testée	CLI v7.0.267	Shell v4.9.20.0	v6.7152209	v1.6.1	v1.8.4	v1.10.4	v7.0.6.215
Testeur	CECO	CECO	CECO	CECO	CECO	CECO	CECO
Date du test	Juin 2017	Juin 2017	Juin 2017	Juin 2017	Août 2017	Décembre 2017	Avril 2018
Remarques	Le consensus et la précision seraient nettement meilleurs avec l'option «*contrôler information N dans OutputIntent*».		Les prix datent de 2010.		Le consensus et la précision se sont massivement détériorés dans la version 1.8.4. 18 des 30 fichiers PDF invalides sont passés pour valides.	La vitesse s'est considérablement améliorée dans la version 1.10.4. Le consensus et la précision sont comparables à la version 1.6.1.	L'utilisation de l'interface de ligne de commande exige une licence de serveur au prix de 1'200 EUR.

Télécharger l'étude (en allemand)

[PDFA-Validatoren_2018.04_v1.0.pdf](#) , 173.7K, 23.04.18
Validateurs PDF/A - version 2018.04

Pour une analyse plus détaillée de PDF/A voir le Catalogue de formats de données d'archivage:

[PDF/A-1](#)
[PDF/A-2](#)
[PDF/A-3](#)

La version 2010 de l'étude sous forme de fichier PDF

[PDFA-Validateurs_v2.1.pdf](#) , 43.3K, 08.12.10
Validateurs PDF/A - version 2010

Contact
A propos
Impressum
Événements
Newsletter
RSS