

Convertir les fichiers PDF en objets éditables

Architectes, Ingénieurs, Entrepreneurs et Industriels, tous les métiers ont besoin de collaborer constamment durant toutes les phases d'un projet – conception, réalisation et opération. Les informations doivent être distribuées à tous les niveaux de la chaîne logistique et les dessins y ont un rôle prépondérant : conception, présentations, schémas, cartes, illustrations, etc...

Clients, fournisseurs et partenaires d'un même projet utilisent bien souvent des applications et plateformes différentes pour générer leurs dessins. Cela conduit bien évidemment à des problèmes de compatibilité et à des pertes de productivité. Manque de contrôle sur les versions et propriété intellectuelle sont deux des raisons pour lesquelles les entreprises rechignent bien souvent à distribuer leurs fichiers de dessin dans un format natif comme le DWG. Le format PDF s'est donc tout naturellement imposé, de par son indépendance vis-à-vis des applications, pour collaborer, distribuer, archiver, échanger et visionner les dessins.

De plus en plus d'entrepreneurs et autres membres de la chaîne logistique reçoivent des dessins au format PDF. Bien souvent l'information contenue dans ces dessins doit être importée dans leur environnement pour des raisons diverses: visionner, tracer, éditer, mesurer, tester, fusionner, etc... Dans bien des cas les demandes de devis et appels d'offre sont accompagnées de fichiers PDF et les entrepreneurs doivent être capables de réutiliser les informations géométriques contenues dans ces dessins afin d'assurer l'exactitude de leur soumission.

La plupart des applications CAO/FAO/CNC ne peuvent pas utiliser le format PDF comme format d'entrée. Et il n'y a pas de façon simple d'extraire la géométrie vectorielle d'un fichier PDF. Bien souvent la seule solution est de redessiner les plans à partir de zéro, ou alors de scanner les copies afin de les insérer dans le logiciel de dessin pour les utiliser comme aide au retraçage.

Peu de gens le savent, mais les fichiers PDF générés par les applications CAO retiennent en fait la géométrie vectorielle des dessins. Un logiciel comme pdf2cad de Visual Integrity pourra en quelques secondes convertir ces PDF vectoriels en fichier DXF importables et éditables dans tout logiciel CAO ou DAO, permettant une incroyable économie de temps et d'énergie.

PDF vectoriel et PDF image

Un fichier PDF peut être soit vectoriel, soit image (ou raster). Pour les différencier, ouvrez votre fichier PDF avec par exemple Acrobat Reader, zoomez à 600-800% sur une partie du dessin. Si les traits sont nets, il s'agit d'un PDF vectoriel. Si les traits sont en fait une suite de petits carrés (pixels) un peu flous, votre fichier est un PDF image ou raster.

Le fichier vectoriel va contenir les informations géométriques du dessin, les PDF images proviennent souvent de scanners et contiennent des images, photos, fichiers TIFF ou JPEG. Les images raster ne sont rien de plus qu'une suite de pixels avec leur couleur propre formant ensemble une représentation graphique du dessin. Il n'y a aucune information géométrique, ou intelligence, dans un PDF image.

La seule solution pour convertir un PDF image en objets éditables est un logiciel de vectorisation. Recsoft EasyDone pour AutoCAD version pleine et Able Software R2V (GIS) sont deux exemples pas trop coûteux, ils existent en version française et cela mérite une mention. Il vous faudra toutefois tout d'abord extraire les images du fichier PDF pour les convertir. Scan2CAD PRO ouvre directement le format PDF, mais il n'est disponible qu'en anglais. Ces programmes tentent de reconnaître les lignes, formes et parfois les textes à partir de suites de pixels. En plus d'être totalement dépendant de la qualité originale des scans, beaucoup de manipulations sont souvent nécessaires afin d'obtenir un bon résultat et presque toujours un gros travail de nettoyage subsiste après la vectorisation.

Si par contre le fichier PDF a été généré par l'application avec laquelle le dessin a été créé (ou avec Acrobat Distiller, Bluebeam, Acroplot ou encore PDF-XChange), le dessin sera vectoriel – une suite d'instructions qui définissent chaque ligne, forme et objet en utilisant des équations mathématiques. En clair une version 2D du dessin, tel qu'il aurait été imprimé sur papier. De plus en plus d'entreprises génèrent maintenant des PDF vectoriels et c'est exactement pour ce genre de fichiers que pdf2cad a été conçu.

Comment fonctionne pdf2cad?

pdf2cad interprète les instructions de dessin contenues dans le fichier PDF et les convertit dans leur équivalent en langage vectoriel DXF. Les types de lignes utilisés dans le PDF sont reproduits dans le DXF. Il est important de préciser que ces types de lignes ne sont pas toujours ce que vous aimeriez qu'ils soient: les fichiers PDF ne reconnaissent pas les cercles ou ellipses, et ne les représente que par des segments de courbes ou polygones. Les courbes peuvent déjà avoir été décomposées en segments de polygone dans le PDF. pdf2cad n'aura pas d'autre choix que de représenter ces types de lignes tels qu'il les trouve dans le PDF. Toutefois Visual Integrity travaille sur de nouvelles fonctions qui seront capables de reconnaître, par exemple, des cercles dans le dessin PDF et de les représenter dans un fichier DXF.

La même histoire est également valable pour les textes. Dans bien des cas, les textes d'un dessin seront déjà convertis en polygones ou splines dans le fichier PDF (par exemple si générés avec AutoCAD). Cela signifie que toute l'information "texte" est perdue et qu'il ne s'agit en fait que de lignes ou splines ayant l'apparence d'un texte. Dans ce cas également pdf2cad ne peut que reproduire cela dans le fichier DXF. Si par contre les textes sont présents dans le fichier PDF, alors pdf2cad sera dans ce cas là en mesure de reproduire ces textes en textes éditables (MTEXT) dans le fichier DXF. Afin de savoir si vos textes seront reconnus par pdf2cad, essayez de les rechercher dans Acrobat Reader. Si vous êtes capables de le faire, alors vos textes seront reconnus comme des textes éditables dans le fichier DXF.

Les images raster contenues dans un fichier PDF sont extraites par pdf2cad dans des fichiers séparés TIFF (noir&blanc) ou JPEG (couleurs), et référencées dans le DXF. Si vous les garder dans le même dossier que votre fichier DXF, elles apparaîtront lors de l'ouverture de ce dernier – pour autant que votre application supporte les images raster. Si le dessin dans votre PDF est une image raster (ou contient des images raster), pdf2cad vous le signifiera dès la conversion terminée.

En résumé, on peut dire qu'un PDF vectoriel n'est pas un fichier original, mais c'est bien plus qu'une simple image.

Notez également que Visual Integrity développe pdf2cad (et ses autres logiciels) en version serveur et SDK pour Intégrateurs et Développeurs. Les versions serveurs permettent l'automatisation du processus de conversion PDF en DXF, accessible via une ligne de commande, un script ou une surveillance de dossier. Les versions SDK permettent d'intégrer la conversion PDF en DXF directement dans une autre application (par exemple logiciel CAO).