

Modifier des fichiers PDF sous Ubuntu

Sylvain Blondeau

Le format de fichier PDF (*Portable Document Format*) fut créé en 1993 par la société Adobe System. Tout d'abord fermé, c'est aujourd'hui un format ouvert disposant même d'une normalisation ISO depuis 2008.

Le but du PDF est de pouvoir conserver la mise en page d'un document (images et graphiques, type, couleur et taille de police...) sur n'importe quel appareil lisant ce fichier. Le format librement utilisable signifie, d'une part, que n'importe quel type de logiciel est en mesure de générer un fichier PDF et, d'autre part, que ce dernier sera lisible à l'identique sur une tierce machine possédant un lecteur de PDF, même si le système d'exploitation est différent.

Remarque : le format PDF continue d'évoluer et gère maintenant des formulaires, c'est-à-dire des champs spécifiques avec lesquels il est possible d'interagir. Il faut cependant que le logiciel de lecture gère ces nouvelles fonctionnalités de PDF. Remplir un formulaire en PDF n'a rien à voir avec une modification de fichier PDF comme nous le décrirons plus loin dans cet article.

Lire un fichier PDF sous Ubuntu

Si Adobe a ouvert le format, la société continue de développer son lecteur de PDF gratuit mais propriétaire, *Acrobat Reader*, largement utilisé sous Microsoft Windows. Sous Ubuntu, plusieurs alternatives libres existent heureusement.

Par défaut, c'est le logiciel Evince qui est utilisé. Il permet, comme la plupart des autres visionneuses de PDF, de zoomer, rechercher du texte, effectuer une rotation des pages, etc. Il existe de nombreuses alternatives à Evince sous GNU/Linux dont, par exemple, Okular (par défaut sous Kubuntu) ou encore xPDF (un lecteur très léger, idéal pour les petites configurations).

Créer un fichier en PDF

S'il est très simple et courant de lire un fichier PDF, il est parfois utile de créer ses propres fichiers. Le logiciel le plus pratique – et qui conviendra dans la majorité des cas – est la suite bureautique LibreOffice (qui intègre cette fonctionnalité depuis très longtemps). Tout fichier créé dans LibreOffice est exportable en PDF, en utilisant l'icône « Export direct au format PDF », juste à côté de l'icône d'impression dans la barre d'outils (Figure 1). Précisons qu'il ne faut pas utiliser *Enregistrer sous* mais bien la fonction *Export*.

Cependant, il est parfois nécessaire de créer un PDF sans recourir à LibreOffice. Ubuntu intègre pour cela une fonction très intéressante, une imprimante PDF. Tout fichier imprimable (quelque soit ce fichier ou le logiciel utilisé) pourra être enregistré en PDF. Lorsque vous essayez d'imprimer un document, une boîte de dialogue s'affiche présentant les imprimantes installées, mais également une ligne *Imprimer dans un fichier*. Sélectionnez cette ligne, choisissez un nom et une destination pour le fichier et choisissez le format PDF (Figure 2). Au lieu d'imprimer le fichier sur du papier via votre imprimante, l'ordinateur utilise cette imprimante virtuelle, troquant ainsi la feuille par un fichier au format PDF. Simple et économique !

Vous êtes maintenant en mesure de créer simplement des fichiers PDF, format idéal pour partager des documents avec vos connaissances,

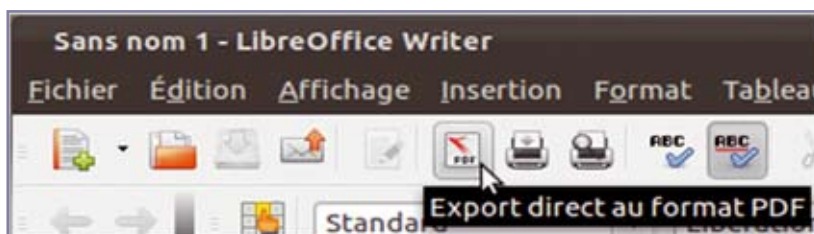


Figure 1. Exporter au format PDF dans LibreOffice

quel que soit leur type d'écran ou de système d'exploitation, puisque la mise en page reste partout la même.

Modifier le contenu d'un PDF

S'il conserve très bien les mises en forme, le format PDF se révèle – car il n'est pas fait pour ça – difficile à modifier. Si la chose était quasiment irréalisable il y a encore quelques années, il devient cependant aujourd'hui de plus en plus facile de modifier tout ou partie d'un document. Nous allons vous présenter trois logiciels permettant de réaliser cette tâche.

LibreOffice Draw : pour les gros fichiers textes

Draw n'est sans doute pas le logiciel le plus utilisé de la suite bureautique LibreOffice, derrière Writer, Calc ou Impress. Cependant, il possède depuis quelques temps une fonctionnalité très intéressante, celle d'éditer les fichiers PDF. L'option n'est pas encore active par défaut ; il faut, pour l'utiliser, installer une extension. Celle-ci est cependant disponible dans les dépôts et donc très facilement installable. Le paquet (disponible dans la logithèque Ubuntu) s'appelle `libreoffice-pdfimport`.

Une fois installée, cette extension permet d'ouvrir un fichier PDF dans LibreOffice en se rendant dans le menu *Fichier* → *Ouvrir* et en faisant attention à bien sélectionner le format PDF en type de fichier.

Nous n'allons pas le cacher, le rendu se trouve alors plutôt malmené, d'autant plus que la mise en forme est complexe et que le nombre d'éléments graphiques est important. En effet, ceux-ci ne sont pas très bien gérés. Cependant, si le PDF contient plusieurs pages de texte, celles-ci s'ouvrent toutes dans le logiciel et il est aisé de modifier alors l'ensemble d'un fichier. Le texte est reconnu sous forme de zones de textes éditables. Néanmoins, cette reconnaissance est très perfectible puisqu'il arrive souvent qu'un même paragraphe soit constitué d'une dizaine de zones de texte différentes au lieu d'en former une seule et même, ce qui risque de compliquer

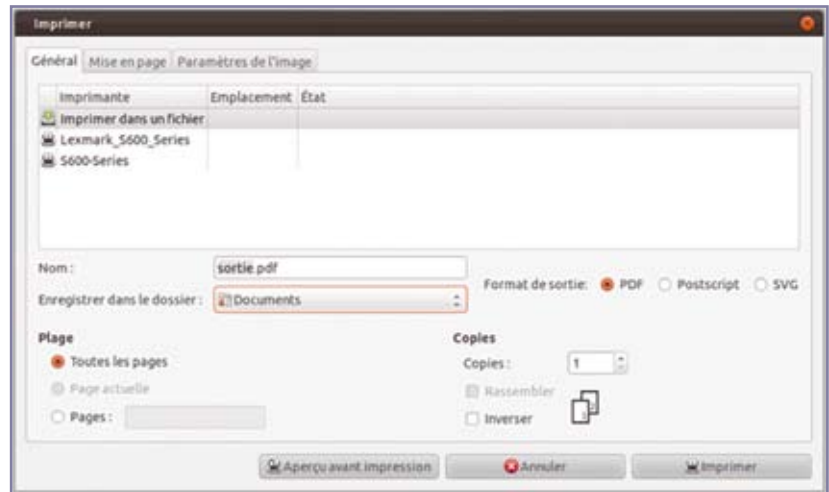


Figure 2. Imprimante virtuelle permettant de créer des fichiers PDF

significativement les modifications. L'utilisation de Draw sera donc à privilégier pour les documents de plusieurs pages, contenant principalement du texte et pour lesquels la conservation de la mise en forme n'est pas une priorité.

Inkscape : pour des modifications précises

Inkscape est un logiciel de dessin vectoriel. Depuis quelques années, il permet d'importer un fichier PDF, il convertit alors ce document au format SVG (donc en dessin vectoriel). Le principal avantage d'Inkscape est qu'il conserve presque parfaitement la mise en page du document. Les éléments graphiques sont justement redessinés, à la bonne place et sous la bonne forme, contrairement à Draw. Il faudra souvent dégrouper plusieurs fois les objets (Inkscape regroupe par défaut certains éléments) afin d'avoir accès à la modification de chaque objet, mais cela est très facile à réaliser. Le texte est, là encore, très bien reconnu (mieux que sous Draw) et forme une zone par paragraphe (même si quelques approximations surviennent parfois). Comme le montre la Figure 3, il est très facile, en utilisant les nombreuses fonctionnalités d'Inkscape, de modifier en profondeur polices, tailles, couleurs ou même, d'ajouter de nouveaux objets graphiques afin de modifier en profondeur une mise en page.

Cependant, le défaut majeur du logiciel tient dans l'impossibilité d'importer un PDF autrement que page par

page (chaque page modifiée pouvant alors être ré-enregistrée dans autant de fichiers PDF indépendants qu'il faudra ensuite fusionner). Cette procédure est assez peu pratique, demande des manipulations supplémentaires et le travail de modification sera à refaire entièrement pour chaque page. L'édition d'un PDF via Inkscape est ainsi rendue assez lente et lourde – même si la prise en main est relativement aisée – d'autant plus qu'il faudra reprendre le travail de mise en page à zéro pour chaque page. Par exemple, modifier la police sur un document texte de cinquante pages devient impossible, alors que la manipulation serait très rapide sous Draw.

Des trois logiciels d'édition présentés – chacun ayant ses spécificités, qualités et défauts – Inkscape reste cependant le plus polyvalent et permet de travailler sur un PDF très fidèle à l'original.

GIMP : à réserver aux images

Le dernier logiciel présenté dans cet article sera le célèbre GIMP, habituellement réservé aux retouches de photos et autres joyeusetés graphiques. Pour faire simple, il est possible d'importer son PDF dans GIMP, ce dernier le transformant tout simplement en une image. Nous noterons que le multi-page est géré en créant un calque par page. À partir d'un PDF, nous obtenons donc une image modifiable à loisir à l'aide de toute la puissance de GIMP : rotations, filtres, découpages, tout est possible (Figure 4). Cependant,

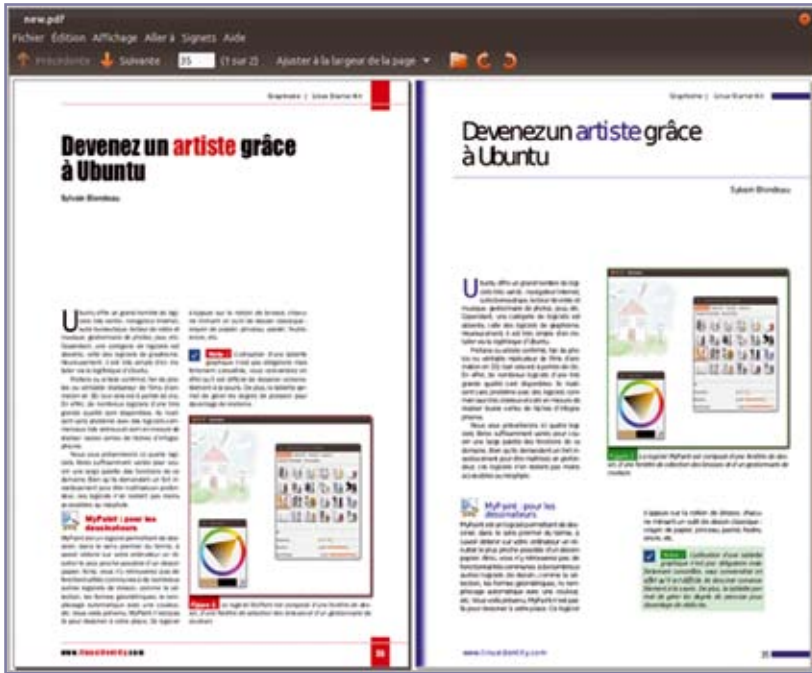


Figure 3. Les possibilités de modification sous Inkscape (à droite) sont nombreuses : polices, couleurs, mise en page...

puisque converti en image, le texte de notre PDF n'est plus reconnu en tant que tel, il devient donc impossible de le modifier. L'intérêt d'utiliser GIMP semble donc limité. Il est cependant

utile si votre fichier PDF n'est en fait qu'une grosse image sans texte et que vous voulez modifier quelque chose sur cette image. Le cas de figure se produisant très souvent est



Figure 4. Avec GIMP, le PDF devient une simple image, modifiable à volonté par le logiciel

celui d'un fichier PDF issu de la numérisation d'un document texte. En effet, si vous scannez une feuille de texte et créez ensuite un PDF, l'ensemble n'est pour l'ordinateur qu'une grande image, le texte n'étant pas reconnu. Sous Draw ou Inkscape, cette image n'est pas vraiment modifiable. Sous GIMP, il sera possible de gommer des éléments et facilement sélectionner des parties pour les déplacer, les redimensionner, les tourner, etc. Bien que peu pratique, cette édition du PDF est tout à fait réalisable dans ce cas.

Découper ou fusionner un PDF

Modifier un PDF ne signifie par forcément en modifier la mise en forme ou son contenu. Il est parfois intéressant de supprimer des pages, de modifier leur ordre ou d'ajouter des pages

provenant d'un autre document PDF. Comme nous l'avons vu précédemment, Inkscape n'édite le document que page par page. Nous obtenons donc au final un PDF découpé dont il faut fusionner les différents éléments. C'est un bon exemple d'utilisation d'un logiciel comme PDF-mod (disponible dans la logithèque Ubuntu) qui permet de réaliser cela très facilement à l'aide d'une interface graphique.

Le fonctionnement de PDF-mod est extrêmement simple : ouvrir un PDF, une vignette par page apparaît alors. Il est possible de modifier l'agencement de ces pages, d'en supprimer ou de les tourner. Il est également possible d'ajouter des pages provenant d'un autre PDF en utilisant le bouton en forme de plus dans la barre d'outil ou par glisser/déposer. C'est ainsi qu'il sera aisé de reconstituer un document multi-page modifié sous Inkscape.

Il existe d'autres logiciels équivalents et permettant même des utilisations plus avancées, dont PDF split and merge ou des utilitaires en ligne de commande comme pdftk.

Astuce : une fois modifié, le poids du fichier PDF est parfois très important. Une astuce (qui fonctionne souvent mais pas à chaque fois car elle dépend du ratio texte/image du document) consiste à utiliser deux courtes lignes de commandes permettant de convertir le PDF en postscript (un format équivalent), puis de le convertir à nouveau en PDF. Pour cela, il faut taper dans un terminal : pdf2ps fichier.pdf fichier.ps puis ps2pdf fichier.ps nouveau_fichier.pdf

Cela prend parfois un peu de temps, mais le nouveau fichier est généralement beaucoup plus léger que l'initial, sans perte de qualité.

Nous venons de survoler brièvement quelques uns des nombreux logiciels gravitant autour du format PDF. C'est un sujet riche et complexe, mais nous retiendrons qu'aujourd'hui, l'environnement GNU/Linux (notamment la distribution Ubuntu) permet de lire, créer et modifier ce type de fichiers de plus en plus facilement, rapidement et efficacement. ■