

Récupérer du texte depuis une image ou un PDF.

Comment récupérer du texte depuis une image ou un PDF ? Cette question, vous vous l'êtes déjà posée. Ça vous est forcément arrivé ! On vous demande de publier un texte sur le site paroissial. Mais plutôt que de vous l'envoyer par mail, on vous a seulement remis une feuille avec le texte imprimé... Comment faire ? La solution : utiliser un OCR.

Qu'est que l'OCR ?

La technologie de reconnaissance optique de caractères (OCR en anglais pour Optical Character Recognition ou ROC en français pour Reconnaissance Optique de Caractères) est l'une des technologies modernes les plus utiles. La technologie OCR vous permet de numériser des documents papier réels et de les transformer en documents numériques lisibles par un ordinateur ou un smartphone en seulement quelques secondes.

source : <https://www.astuces-aide-informatique.info/6561/les-5-meilleurs-logiciels-ocr-gratuits>

Des outils gratuits en ligne

source : <https://www.futura-sciences.com/tech/questions-reponses/outils-pdf-logiciels-ocr-sont-meilleurs-11568/>



LightPDF est un service OCR en ligne qui permet de convertir en quelques clics un document ou une image en un fichier éditable.

Sa formule OCR Basic est entièrement gratuite. Disposant d'une interface d'une extrême simplicité, l'utilisateur n'a qu'à sélectionner son document PDF ou son image au format JPG ou PNG, choisir le langage employé parmi 25 disponibles et cliquer sur un bouton pour le convertir au format TXT.

<https://lightpdf.com/>



Online OCR est un service en ligne entièrement gratuit qui permet de convertir jusqu'à 15 fichiers par heure en mode Invité. Si l'utilisateur s'enregistre, il a la possibilité de convertir des documents PDF comportant plusieurs pages (contre une seule normalement) et a accès à des fonctions supplémentaires.

<https://www.onlineocr.net/fr/>

Free Online OCR

New OCR est un service en ligne 100 % gratuit qui permet de convertir de nombreux formats de documents tels que PDF, JPG, PNG, GIF, BMP, TIF et Djvu (format d'archives). En sortie, l'utilisateur a le choix entre 3 formats : PDF, DOC et TXT.

Sa grande force est de reconnaître 150 langages, dont certains introuvables sur d'autres outils en ligne (vieux français, vieil anglais, breton, occitan, grec ancien, équation mathématique, etc.). Il a aussi la possibilité d'employer directement Google Translate et Bing Translator s'il a besoin de traduire son document, mais également de l'éditer en ligne avec Google Doc.

<https://www.newocr.com/>

4 mars 2021