



FICHE TECHNIQUE · CRÉER · AI-OCR

Numérisé, mais impossible à rechercher ?

Désormais, chaque scan devient du texte.

AI-OCR lit les scans, photos et PDF et fournit le texte avec la position et l'indice de confiance de chaque ligne, ainsi que les codes QR et codes-barres, les tableaux et un PDF consultable. Pour les équipes qui travaillent avec du papier et des images : dans l'application, via l'API REST et via MCP.

EN UN COUP D'ŒIL

Trois niveaux : Basic lit le texte avec les positions, Premium ajoute les codes QR et codes-barres ainsi qu'un PDF consultable, Ultra fournit en plus les tableaux, la mise en page et l'écriture manuscrite.

01 · DE QUOI IL S'AGIT

Un scan reste muet pour la recherche. AI-OCR le rend lisible.

Pour toute recherche, une page numérisée n'est qu'une image. Si l'on a besoin de son contenu, il faut le retaper ou la feuilleter à la main. AI-OCR reconnaît le texte de chaque page, mémorise où se trouve chaque ligne et restitue le tout sous forme de texte, de Markdown, de JSON ou de PDF consultable.

ENTRÉES

PNG, JPG, TIFF, BMP, WEBP et PDF, y compris multipages ; dans l'application jusqu'à 20 Mo par fichier

RÉSULTAT

Texte, Markdown et JSON avec cadre englobant et confiance par ligne, plus un PDF consultable

ACCÈS

Dans l'application, via l'API REST et comme outil MCP `po_aiocr_analyze`

LANGUE

La langue de chaque page est détectée automatiquement et figure dans le résultat

**Texte avec position**

Chaque ligne reconnue est fournie avec un cadre englobant et un indice de confiance. Vous pouvez ainsi vérifier et marquer chaque résultat sur la page.

Codes QR et codes-barres

À partir de Premium, AI-OCR lit aussi les codes QR et codes-barres de la page et restitue leur contenu avec leur position dans le résultat.

PDF consultable

À partir de Premium, un PDF sandwich est créé : le scan d'origine avec une couche de texte invisible, consultable et copiable.

Tableaux en données

Ultra détecte les tableaux avec lignes et colonnes et les restitue sous forme de données structurées en JSON, avec la mise en page de la page.

Écriture manuscrite

Ultra reconnaît aussi l'écriture manuscrite, par exemple les notes sur les formulaires et les bons de livraison, et la restitue avec sa position, comme le texte imprimé.

Partie intégrante de l'AI-DMS

Dans le workspace, AI-OCR fonctionne dans le cadre du traitement AI-DMS. Ensuite, le texte des pages traitées est disponible dans la recherche.

02 · FONCTIONNEMENT

Importez un fichier, choisissez un niveau. Vérifiez le résultat et utilisez-le.

Dans l'application, vous importez un fichier et voyez le résultat directement sur la page. Via l'API REST et MCP, la même tâche s'exécute avec les mêmes trois niveaux.

1

Importer un fichier

Faites glisser une image ou un PDF dans le champ ou cliquez pour le choisir : PNG, JPG, TIFF, BMP, WEBP ou PDF jusqu'à 20 Mo.

2

Choisir le niveau et lancer

Sous **Choisir le mode OCR**, choisissez Basic, Premium ou Ultra, puis cliquez sur **Démarrer AI-OCR**. Les PDF multipages sont lus page par page.

3

Vérifier le résultat

Dans l'onglet **Positions du texte**, cliquez sur une zone : le texte, la position et la confiance s'affichent à côté. **Tableaux** présente les tableaux détectés.

4

Transmettre

Poursuivez avec **Télécharger le PDF recherchable** ou **Exporter JSON**. Via l'API : POST /job/add/paperoffice_aiocr___generate avec ocr_mode.

CARACTÉRISTIQUES TECHNIQUES

FONCTIONNALITÉ	DÉTAILS
Formats d'entrée	PNG, JPG, TIFF (y compris multipages), BMP, WEBP, PDF (multipages)
Sortie	Texte, Markdown, JSON avec cadres englobants et confiance ; tableaux et mise en page en JSON ; PDF sandwich
Niveaux (ocr_mode)	basic = text, premium = grid, ultra = complete
API REST	POST /job/add/paperoffice_aiocr___generate, résultat dans la réponse ou par identifiant de tâche
Outil MCP	po_aiocr_analyze, fichier via upload_id, file_url ou pofid du workspace
Dans le workspace	Lors du traitement AI-DMS, lancement via Actions › Agents et workflows › AI-DMS
Langue	Détection automatique par page

03 · EN PRATIQUE

Du papier devient des données. Sans retaper, sans modèle.

Rendre une ancienne archive consultable

Les dossiers numérisés passent par Premium et reviennent sous forme de PDF sandwich. Ensuite, vous pouvez rechercher, marquer et copier le texte de chaque PDF.

Positions issues des bons de livraison

Ultra détecte le tableau des positions d'un bon de livraison. Les lignes sont fournies sous forme de tableau structuré en JSON et peuvent être traitées sans être retapées.

Texte reconnu dans votre application

Une équipe de développement envoie des photos de sa propre application à l'API REST et reçoit en retour du texte avec des positions, par exemple pour marquer des champs directement sur l'image.

QUESTIONS FRÉQUENTES

Quelle est la différence entre Basic, Premium et Ultra ?

Basic fournit le texte, le Markdown et les cadres englobants avec confiance. Premium ajoute les codes QR et codes-barres ainsi qu'un PDF consultable. Ultra fournit en plus les tableaux, la mise en page avec ordre de lecture et l'écriture manuscrite.

AI-OCR peut-il lire les documents de mon workspace ?

Oui. L'outil MCP `po_aiocr_analyze` accepte le `pofid` d'un document enregistré, par exemple dans Claude ou Cursor. Lors du traitement AI-DMS dans le workspace, AI-OCR s'exécute de toute façon.

Dois-je indiquer la langue du document ?

Non. AI-OCR détecte automatiquement la langue de chaque page et la restitue dans le résultat.

Comment savoir si la reconnaissance est fiable ?

Chaque ligne porte un indice de confiance, et l'application indique sa position sur la page. Cette valeur se réfère à l'échelle de référence de PaperOffice AI.

COMMENCER

Essayez avec votre scan le plus difficile.

Connectez-vous à votre compte PaperOffice, ouvrez AI-OCR, chargez un fichier et cliquez sur **Démarrer AI-OCR**. La facturation se fait à l'usage et dépend du niveau.

paperoffice.ai/fr/ai-ocr-reconnaissance-texte-llm-numerisation



EN SAVOIR PLUS SUR LE MODULE

paperoffice.ai/fr/ai-ocr-reconnaissance-texte-llm-numerisation

TARIFS

paperoffice.ai/fr/tarifs

CONTACT

paperoffice.ai/fr/contact