

Livre blanc Business

Transformez vos documents numérisés en textes interrogeables et modifiables

avec la technologie OCR d'I.R.I.S[®] intégrée

Imprimantes multifonctions et Digital Sender HP Flow



Sommaire

4 Les bases

4 Qu'est-ce que la numérisation vers OCR intégrée ?

4 Comment l'OCR fonctionne-t-elle ?

4 Comment obtenir les meilleurs résultats ?

5 Technologie OCR d'I.R.I.S.

5 Formats de fichier de sortie pris en charge

5 Langues prises en charge

5 Comment effectuer une numérisation vers OCR ?

6 Numérisation vers OCR depuis le panneau de commande

7 Créer un Quick Set personnalisé de numérisation vers OCR

11 Numériser en OCR avec un Quick Set personnalisé

12 Comment préciser la langue OCR

13 Meilleures pratiques : conseils pour bien utiliser l'OCR

13 Ce qui peut et ne peut pas être lu par l'OCR

13 Suggestion de paramètres pour obtenir de meilleurs résultats

13 Utilisation de la prévisualisation d'images pour tester et optimiser les numérisations

14 Tailles de fichier et temps de traitement OCR

14 Exemples de numérisation

14 Conditions des tests de numérisation

15 À propos du logiciel HP Digital Sending en option

16 En conclusion

Exploitez au mieux vos données grâce à l'OCR intégrée et aux nombreuses options d'envoi de vos numérisations

Donnez des moyens à vos équipes et aidez-les à être plus productives avec les imprimantes multifonctions et Digital Senders HP Flow, dotées de nombreuses fonctionnalités. Avec un chargeur automatique, la numérisation recto-verso en un seul passage, la détection de double alimentation par ultra-sons et le traitement embarqué de la reconnaissance optique des caractères (OCR)^{1,2}, ces appareils sont un gage de performance pour le traitement de votre flux de travail³.

Ce livre blanc a pour but de vous aider à comprendre les capacités et les limites de la technologie OCR elle-même et du traitement OCR sur l'appareil. Ces informations peuvent vous guider pour intégrer les capacités de l'OCR à votre flux de travail.

Les bases

Qu'est-ce que la numérisation vers OCR intégrée ?

La reconnaissance optique des caractères (OCR) vous permet de convertir des images numérisées dans des formats de fichier courants avec du texte interrogeable et modifiable. Les fonctions OCR facilitent également l'indexage et l'extraction des documents. Quand vous utilisez le scanner intégré, tous les processus OCR (fonction de reconnaissance optique des caractères) s'effectuent au niveau de l'appareil, et non plus sur un ordinateur ou un serveur hôte, comme dans le cas des logiciels OCR¹.

Comment l'OCR fonctionne-t-elle ?

Le logiciel OCR analyse et transforme les formes trouvées dans l'image numérique d'une page de texte en caractères, puis les enregistre dans un format permettant d'indexer le document ou d'y faire des recherches sur un ordinateur : PDF interrogeable, RTF (format texte enrichi), Unicode ou texte ASCII.

Comment obtenir les meilleurs résultats ?

Le logiciel s'efforce de reconnaître les formes des caractères. Il est donc important que le document numérisé soit aussi net (contraste élevé) et propre que possible. Plusieurs facteurs supplémentaires contribuent à la réussite d'une lecture précise. Nous les évoquerons plus en détails ultérieurement (voir « Meilleures pratiques : conseils pour bien utiliser l'OCR », page 11).

¹ Pour accéder à la liste des appareils dotés des fonctionnalités OCR, consultez le guide *HP Scanning Devices Selection Guide*, disponible en anglais à l'adresse <http://h20195.www2.hp.com/v2/GetDocument.aspx?docname=4AA4-3848EEW>.

² Sur les imprimantes multifonctions et Digital Sender HP Flow, l'OCR intégrée est destinée à une utilisation occasionnelle. Pour une utilisation intensive de l'OCR, il est conseillé d'utiliser le logiciel DSS (Digital Sending Software), basé sur un serveur, proposé en option.

³ D'autres imprimantes multifonctions et d'autres scanners peuvent également numériser vers OCR mais nécessitent un logiciel en option tel que DSS (HP Digital Sending Software). HP Digital Sending Software est une option vendue séparément. Voir page 13 pour plus d'informations.

Technologie OCR d'I.R.I.S.

Les imprimantes multifonctions et Digital Sender HP Flow multifonctions utilisent la technologie OCR intégrée d'I.R.I.S., spécialiste de longue date de l'OCR. L'OCR vous permet de numériser des documents papier et de les convertir en PDF interrogeables et en fichiers texte totalement modifiables. C'est une fonction essentielle des flux de travail les plus avancés, car elle permet de procéder à l'indexation des documents et d'y faire des recherches lors de l'envoi vers différentes destinations, y compris les dossiers en réseau, les e-mails, SharePoint® ou une mémoire flash USB.

Formats de fichier de sortie pris en charge

Les imprimantes multifonctions prennent en charge plusieurs formats de sortie pour l'OCR.

- **PDF interrogeable (OCR)** : crée un fichier PDF avec du texte que vous pouvez sélectionner ou dans lequel vous pouvez faire une recherche, tout en préservant l'apparence du document numérisé.
- **PDF/A (OCR)** : un format PDF conçu pour l'archivage à long terme de documents numériques. Toutes les informations sur le formatage du document sont intégrées à ce format.
- **RTF (OCR)** : crée un fichier au format RTF (Rich Text). Le RTF est un autre format de texte qui peut être ouvert par la plupart des programmes de traitement de texte (compatibles avec Microsoft® Word). Certains des éléments de la mise en page initiale seront sauvegardés en utilisant cette option.
- **Texte (OCR)** : crée un fichier texte ASCII (TXT) qui peut être ouvert par tout programme de traitement de texte. ASCII offre une prise en charge limitée des alphabets élargis, utilisés par de nombreuses langues non anglaises. Avec cette option, la mise en page initiale n'est pas sauvegardée.
- **Texte Unicode (OCR)** : une norme professionnelle utilisée pour représenter de façon constante un texte dans n'importe quelle langue. Les langues qui utilisent des caractères non romains doivent utiliser Unicode pour les fichiers TXT. Avec cette option, la mise en page initiale n'est pas sauvegardée.
- **CSV (OCR)** : utilise le format de valeur séparée par une virgule (CSV). Ce type de fichier est recommandé pour la lecture de feuilles de calcul et peut être ouvert par la plupart des tableurs traitant les mots ou des programmes de bases de données.
- **HTML (OCR)** : crée un fichier hypertext markup language (HTML). HTML sert à publier les fichiers sur les sites internet. (Si le document initial contient des images, le résultat obtenu sera stocké dans un fichier zip.)

Langues prises en charge

Le logiciel OCR intégré peut lire les 27 langues suivantes : allemand, anglais, catalan, chinois (simplifié), chinois (traditionnel), coréen, croate, danois, espagnol, finnois, français, grec, hongrois, indonésien, italien, japonais, néerlandais, norvégien, polonais, portugais, roumain, russe, slovaque, slovène, suédois, tchèque et turc.

Vous pouvez modifier les langues de l'OCR depuis le panneau de commande (les instructions débutent à la page 10).

Comment effectuer une numérisation vers OCR ?

Il y a plusieurs façons de numériser des documents pour une utilisation de l'OCR. Vous pouvez tout faire depuis le panneau de commande sur l'écran couleur tactile (voir les instructions à partir de la page 5). Vous pouvez également utiliser un HP Quick Set personnalisé pour lancer un flux de travail OCR avec les paramètres que vous avez choisis. Avec les Quick Sets, les utilisateurs peuvent trouver immédiatement ce dont ils ont besoin, sans passer du temps devant le panneau de commande de l'appareil à la recherche des réglages voulus. Voir page 6 les instructions Quick Set.

⁴ Pour accéder à la liste des appareils dotés des fonctionnalités OCR, consultez le guide *HP Scanning Devices Selection Guide*, disponible en anglais à l'adresse <http://h20195.www2.hp.com/v2/GetDocument.aspx?docname=4AA4-3848EEW>.

Numérisation vers OCR depuis le panneau de commande

Pour numériser depuis le panneau de commande sans utiliser les Quick Sets personnalisés :

Modifier les paramètres par défaut depuis le panneau de commande

Vous aurez peut-être besoin de changer les paramètres de numérisation par défaut depuis le panneau de commandes de l'imprimante multifonctions. Vous pouvez le faire en passant par le menu **Administration, Scan/Digital Send Settings**.



1. Placez le(s) document(s) à numériser dans le chargeur automatique ou sur la vitre du scanner et choisissez l'emplacement où vous souhaitez enregistrer le fichier numérisé (dans cet exemple, une clé USB). Saisissez un nom de fichier pour votre fichier enregistré et appuyez sur **File Type**. Un menu déroulant s'affiche.



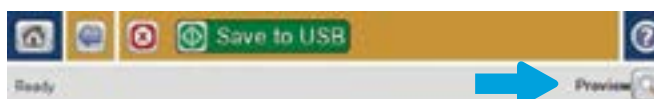
2. Faites dérouler la liste et choisissez le **type de fichier OCR** qui correspond le mieux à votre besoin.



3. Si vous souhaitez personnaliser des paramètres de numérisation, appuyez sur **More Options** pour définir un grand nombre de paramètres, comme le recadrage, la suppression des pages blanches, la qualité et plus encore.



4. Appuyez sur le bouton vert **Save to...**. L'imprimante multifonctions numérise et traite le document. (Vous pouvez choisir **Preview** pour prévisualiser la numérisation dans un premier temps. Voir « Utilisation de la prévisualisation d'image pour tester et optimiser les numérisations », page 11).



5. Le panneau de commande vous informe en permanence sur l'avancement de la tâche. Lorsque la numérisation est terminée, un message d'état ou **Status** indique **Success** (en cas d'échec, il vous informera du problème).



Trouver l'adresse IP de votre appareil

Imprimantes multifonctions : pour afficher l'adresse IP sur le panneau de commande, appuyez sur l'icône d'info sur la connexion.

Scanners : Appuyez sur le bouton **Administration**, puis sur **Reports** (Scanjet Enterprise 8500 fn1) ou **Information** (Scanjet 7000n/nx).

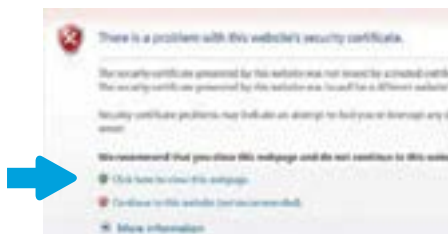


Remarque

L'administrateur peut désactiver le bouton Info sur le panneau de commande. Si le bouton n'apparaît pas, vous pouvez trouver l'adresse IP dans le menu Administration (Administration > Reports > Configuration/Status Pages > How to Connect Page). Si celui-ci n'est pas non plus accessible, contactez votre administrateur pour lui demander l'adresse IP.

Erreur de certificat ?

Si vous voyez une erreur de certificat de sécurité, cliquez sur « **Continue to this website (not recommended).** » (Ne vous inquiétez pas, c'est parfaitement sûr !)



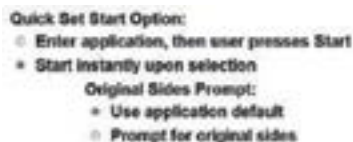
Emplacement et option du bouton Quick Set

• **Emplacement du bouton.** Ce menu déroulant vous permet de définir si votre nouveau bouton Quick Set figurera en tant que nouvel élément sur l'écran d'accueil ou parmi les boutons Quick Set.



Options de démarrage Quick Set

• **Enter application, then user presses Start :** l'appareil se met en pause et attend vos instructions concernant l'opération de numérisation. **Start instantly upon selection :** numérise les documents, qu'ils soient dans le chargeur automatique ou sur la vitre, dès que vous appuyez sur le bouton Quick Set (sauf si vous avez activé l'option Prompt for original sides).



Créer un HP Quick Set personnalisé de numérisation vers OCR

Il est possible de créer des HP Quick Sets personnalisés en utilisant le serveur Internet intégré (EWS) de l'appareil. Dans cet exemple, nous allons créer un Quick Set Enregistrer sous USB, mais vous pouvez souhaiter envoyer votre sortie vers un dossier en réseau, le destinataire d'un e-mail ou une autre destination, comme SharePoint. (Notez que tout réglage Quick Set peut être remplacé sur le panneau de commande ou modifié ultérieurement dans l'EWS pour optimiser les résultats OCR, si nécessaire.)

1. Avec un navigateur, tel que Microsoft Internet Explorer, ouvrez l'EWS en saisissant l'adresse IP de l'appareil dans la barre d'adresse (voir note à gauche).
2. La page d'accueil s'ouvre. Cliquez sur **Sign in** dans le coin supérieur droit et saisissez vos identifiants d'administrateur (compte d'administrateur et mot de passe).



3. Cliquez sur l'onglet **Scan/Digital Send** (A). Sélectionnez, dans le menu de navigation de gauche, la destination désirée pour les fichiers issus de la numérisation (B). Si ce n'est pas encore fait, cochez la case **Enable Save to...** (C) ; vous pouvez aussi choisir d'attendre pour activer, puis cliquer sur **Add** (D).



4. Une boîte de dialogue **Quick Set Wizard** s'ouvre. Donnez une description et un titre au Quick Set OCR, sélectionnez **Button Location** pour l'emplacement du Quick Set et **Start Option** pour l'interaction avec l'utilisateur sur le panneau de commande, puis cliquez sur **Next**.



Création instantanée de dossier

Si vous précisez un chemin ou un dossier qui n'existe pas encore, la première fois que vous utilisez Quick Set, l'imprimante multifonctions vous demandera si vous souhaitez le créer automatiquement.



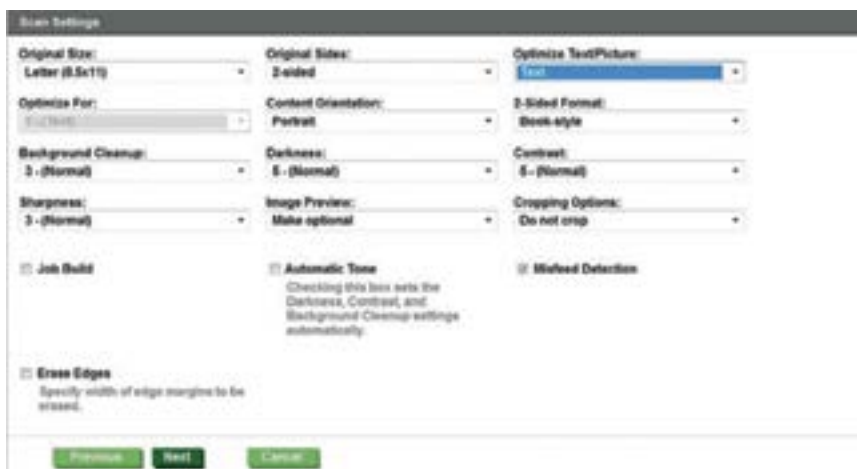
5. L'écran suivant présente les options d'emplacement pour l'enregistrement des fichiers. Dans cet exemple, nous indiquons un dossier nommé « Brokerage Apps » dans le dossier « Confidential ». (Dans l'assistant Quick Set pour l'envoi vers d'autres emplacements, comme des dossiers en réseau, une série d'options différente est proposée selon les paramètres des dossiers.)



6. L'écran suivant vous permet de définir un paramètre de notification pour choisir d'adresser un e-mail à l'utilisateur lorsque la tâche est terminée, ou uniquement en cas d'échec de la tâche.



7. Choisissez tous les autres **Scan Settings** (paramètres de numérisation) souhaités, puis cliquez sur **Next**.



En savoir plus sur les paramètres de fichiers

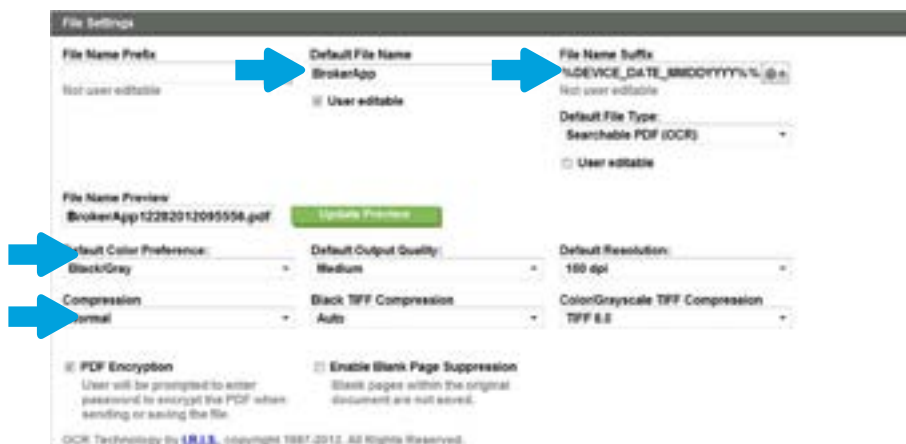
- **Les options File Name Prefix/Suffix** (préfixe/suffixe du nom de fichier) permettent d'organiser les fichiers en définissant les informations qui apparaissent au début ou à la fin du nom de fichier. Utilisez **File Name Preview** pour vérifier les résultats.
- **Default Resolution** permet de définir la résolution de **sortie** de votre scanner en nombre de points par pouce (ppp). Vous pouvez remplacer le réglage par défaut depuis le panneau de commande au moment de la numérisation. D'une manière générale, plus la résolution est élevée, plus le fichier sera lourd.

REMARQUE

Pour les numérisations OCR, les multifonctions **capturent** toujours l'image à 300 ppp, quel que soit le paramètre de **sortie**. Par conséquent, le fait d'augmenter ce paramètre n'améliore pas la performance de l'OCR. Néanmoins, des niveaux de résolution plus élevés peuvent améliorer la qualité des graphismes.

- **PDF Encryption** : utilisez cette option si vous avez besoin d'un fichier PDF sûr et chiffré. Vous serez invité à créer un mot de passe avant l'envoi, dont vous aurez besoin pour ouvrir le fichier.
- **Enable Blank Page Suppression** : il s'agit d'une option utile pour les documents numérisés qui comportent des pages blanches. Les fichiers PDF en particulier seront plus légers du fait de la suppression des pages blanches.

8. Choisissez les **File Settings** (paramètres de fichier) souhaités. Dans cet exemple, « BrokerApp » a été choisi comme **Default File Name** (nom de fichier par défaut), avec un **File Name Suffix** (suffixe de nom de fichier) contenant la date et l'heure. Nous définissons le **Default File Type** (type de fichier par défaut) comme **Searchable PDF (OCR)** et le **Default Color Preference** (préférence de couleur par défaut) comme niveaux de gris car nous n'avons pas besoin de numérisations en couleur. Nous avons laissé le niveau de résolution par défaut à 150 dpi.



9. Vérifiez la boîte de dialogue **Summary** et, si tout semble correct, cliquez sur **Finish**.



10. Un message de confirmation apparaît, confirmant la création du Quick Step ou de toute autre action demandée. Cliquez sur **Apply** et, si vous avez coché la case **Enable Save to...** à l'étape 3 de la page 6, votre nouveau Quick Set apparaît maintenant sur le panneau de commande de l'appareil, prêt à l'emploi.



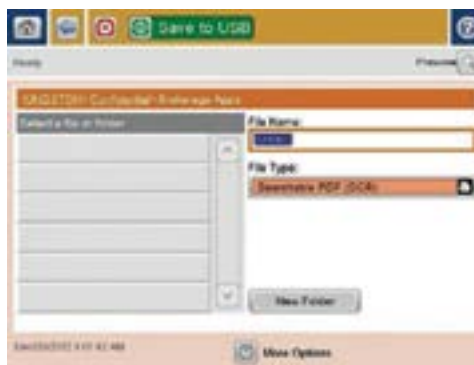
Numériser en OCR avec un Quick Set personnalisé

Lorsque vous avez défini un Quick Set OCR personnalisé et appliqué les modifications depuis l'EWS, celui-ci apparaît sur le panneau de commande de l'appareil.

1. Si vous avez choisi un emplacement de bouton dans « Application Quick Sets » à l'étape 4 de la page 6, appuyez sur le bouton **Quick Sets** pour trouver votre Quick Set. (Si vous avez choisi « Home Screen », votre nouveau Quick Set apparaîtra sur l'écran d'accueil.)



2. Placez le(s) document(s) que vous voulez numériser dans le bac du chargeur automatique ou sur la vitre du scanner à plat, puis appuyez sur votre Quick Set. Selon les paramètres que vous aurez spécifiés, votre tâche de numérisation peut être traitée immédiatement ou, si vous avez choisi l'option de démarrage Quick Set **Enter application then user presses start** (étape 4 de la page 6), vous verrez un écran comme celui figurant ci-dessous. Une fois prêt, appuyez sur le bouton vert **Save to....** L'imprimante multifonction numérise et traite le document. (Vous pouvez choisir Preview pour prévisualiser la numérisation dans un premier temps. Voir « Utilisation de la prévisualisation d'images pour tester et optimiser les numérisations », page 11).



3. Le panneau de commande vous informe en permanence sur l'avancement de la tâche. Lorsque la numérisation est terminée, un message d'état **Status** indique **Success** (en cas d'échec, il vous informera du problème).



En savoir plus sur les paramètres de langue OCR

Au moment de la rédaction de ce document, le seul moyen de définir la langue OCR est d'utiliser le panneau de commande de l'appareil.

Avec les futures mises à jour du micrologiciel, il sera peut-être possible de définir la langue OCR par un Quick Set en utilisant le serveur internet intégré (EWS).

Veuillez noter qu'on peut également modifier la langue OCR par défaut pour **tous les fichiers numérisés** vers une destination donnée, de la façon suivante.

Pour modifier les paramètres de langue OCR par défaut :

1. Depuis le panneau de commande, sélectionnez le bouton du menu **Administration** (un code d'accès peut être demandé).
2. Sélectionnez **Scan/Digital Send Settings**.
3. Sélectionnez la destination du fichier numérisé dont vous souhaitez modifier la langue par défaut.



4. Sélectionnez **Default Job Options** puis **Document File Type** et appuyez sur **File Options**, enfin choisissez la langue de votre choix. Tous les fichiers numérisés OCR liront désormais cette langue par défaut.



Comment préciser la langue OCR

Si vous avez choisi l'option **Enter application then user presses start** (voir étape 4 de la page 6), vous pouvez facilement préciser la langue OCR avant la numérisation du document.

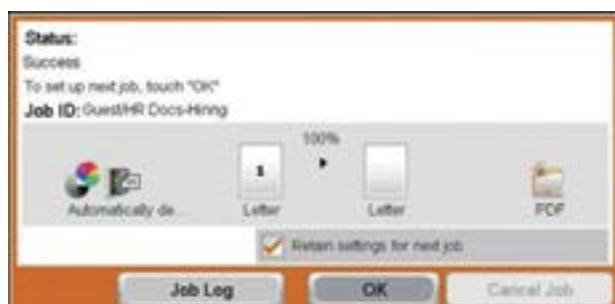
1. Placez le(s) document(s) à numériser dans le chargeur automatique (ou sur la vitre du scanner) et choisissez l'emplacement où vous souhaitez enregistrer le fichier numérisé (dossier en réseau, USB, SharePoint, etc.). Saisissez un nom de fichier pour votre fichier enregistré et appuyez sur **File Type** (même si le type de fichier est déjà correct, cela vous permettra de modifier les options de fichier). Un menu déroulant s'affiche.



2. Appuyez sur le bouton vert **Save to...**. L'appareil numérise et traite le document. (Vous pouvez choisir **Preview** pour prévisualiser la numérisation dans un premier temps. Voir « Utilisation de la prévisualisation d'images pour tester et optimiser les numérisations », page 11).



3. Lorsque la numérisation est terminée, vous pouvez sélectionner **Retain settings for next job**. En cochant cette case, vous numériserez immédiatement d'autres documents avec les mêmes paramètres de numérisation et la même destination.



Meilleures pratiques : conseils pour bien utiliser l'OCR

Ce qui peut et ne peut pas être lu par l'OCR

La technologie OCR a fait des progrès considérables ces dernières années. Cependant, la réussite des numérisations OCR dépend de plusieurs facteurs, notamment des propriétés du document initial. Par exemple, lorsque l'original est net et propre, sans filigrane ou élément perturbateur, le texte obtenu peut être d'une précision très élevée. Si vous n'êtes pas certain que les documents source conviennent, vous pouvez effectuer un test pour évaluer l'équilibre entre les options de saisie manuelle et de reconnaissance optique. Cela peut vous permettre d'établir une stratégie OCR correspondant à vos besoins pour vos flux de documents.

Voici quelques exemples de documents numérisés qui peuvent permettre d'obtenir des résultats très précis.

- Texte net et clair sur une page blanche dans une police de caractères usuelle non décorative
- Documents sans fond de couleur, filigrane contrastant ou autre élément perturbateur

Il est déconseillé d'utiliser l'OCR avec certains documents, par exemple :

- Codes barres
- Textes écrits à la main ou saisis dans un style imitant l'écriture manuelle
- Police de caractère décorative ou ancienne, non usuelle
- Documents à faible contraste, documents avec une impression décolorée ou imprimés sur papier coloré
- Documents ayant été pliés (le pli parcourant le texte peut provoquer des erreurs de lecture)
- Polices de caractère inférieures à 8 points

Suggestion de paramètres pour obtenir de meilleurs résultats

- Si certains de vos documents comportent des pages blanches (par exemple, document recto-verso avec un nombre de pages impair se terminant par une page blanche), activez **Blank Page Suppression** pour réduire au minimum le poids du fichier. Ceci est particulièrement important pour la numérisation de fichiers PDF.
- **Crop to content** permet également de réduire le poids des fichiers PDF ainsi que leur temps de traitement.
- Si la couleur n'est pas utile, réglez le scanner sur noir ou **niveaux de gris**. Ceci peut réduire le poids des fichiers ainsi que le temps de traitement.

Utilisation de la prévisualisation d'images pour tester et optimiser les numérisations

Le panneau de commande de ces appareils dispose de l'aperçu de l'image, ce qui vous permet de diminuer le nombre d'étapes et d'erreurs. Vous pouvez visualiser et modifier vos numérisations directement sur le panneau de commande à écran couleur tactile de l'appareil, avant de les envoyer vers un dossier en réseau, un télécopieur, une adresse e-mail ou un périphérique USB. Vous pouvez facilement obtenir des documents numérisés ayant l'apparence que vous désirez. Avoir immédiatement un aperçu de vos numérisations et pouvoir les rectifier avant leur envoi dans le flux rend les allers-retours avec un ordinateur inutiles.

Numérisez plusieurs pages avec le chargeur automatique de documents et visualisez-les sous forme de vignettes ; ou agrandissez-les et utilisez le mode panoramique pour un examen plus approfondi. Avant de finaliser un document, vous pouvez ajouter une page au fichier numérisé, remettre en ordre les pages, faire pivoter une ou plusieurs pages ou supprimer des pages non désirées. Vous pouvez appliquer des modifications sur une seule page ou sur plusieurs pages simultanément.

Ajouter des page(s)

Remettre en ordre les page(s)

Faire pivoter à 90°

Supprimer des page(s)



Limites des tests informels

Veillez noter que ces exemples de résultats de numérisation sont conçus pour donner une image approximative des poids de fichier et des temps de traitement pour des exemples généraux de types de document. Ces données peuvent vous aider à planifier votre charge de travail selon vos exigences en termes de stockage de fichiers et de temps, mais ne sont pas conçues comme un guide définitif.

Les résultats réels peuvent varier en fonction d'un certain nombre de facteurs, notamment le document scanné et l'appareil de numérisation, ainsi que la version du micrologiciel utilisé (voir Conditions des tests de numérisation ci-dessous). HP développe en continu le micrologiciel pour ajouter des fonctionnalités et optimiser l'équilibre entre performances, taille de fichier et qualité d'image.

Tailles de fichier et temps de traitement OCR

Numériser avec OCR prend plus de temps que la numérisation sans OCR.

Exemples de numérisation

Le tableau ci-dessous vous donne une comparaison des temps de traitement et des poids de fichiers pour les différents types de fichiers OCR, ainsi que deux exemples de numérisation de documents non OCR. Ces tests de numérisation ont été réalisés sur une HP LaserJet Enterprise flow MFP M525c. Les scanners en réseau, les futures versions de micrologiciel et les futures imprimantes multifonctions flow produiront des résultats différents.

- **Texte seul** : document de cinq pages sans graphisme ni photo, en police Arial corps 10 (sans-serif).
- **Texte et photos** : cinq pages de textes avec des puces, un tableau et quelques photos de taille moyenne.
- **Texte et graphismes** : cinq pages de texte avec des graphismes, extraites de ce livre blanc.
- **Feuille de calcul** : une page de feuille de calcul avec 17 colonnes et 25 lignes de données en police corps 8.

Types de document	Formats de fichiers OCR							Non OCR	
	PDF interrogeable	PDF/A interrogeable	RTF	Texte	Texte Unicode	HTML ⁵	CSV	PDF	XPS
Texte uniquement (5 pages, noir) Taille du fichier (Ko)	230	190	27	17	33	38	n/a	103	196
Temps de traitement (min:sec)	1:59	2:05	1:27	1:21	1:18	1:29	n/a	0:11	0:12
Texte et photos (5 pages, noir) Taille du fichier (Ko)	1016	1590	165	18	36	183	n/a	151	257
Temps de traitement (min:sec)	2:41	2:59	2:06	1:28	1:28	2:00	n/a	0:16	0:16
Texte et photos (5 pages, couleur) Taille du fichier (Ko)	1144	1744	164	18	18	224	n/a	1515	1540
Temps de traitement (min:sec)	2:57	3:07	2:26	1:52	1:46	2:21	n/a	0:13	0:16
Texte et graphismes (5 pages, noir) Taille du fichier (Ko)	556	887	355	12	23	404	n/a	713	769
Temps de traitement (min:sec)	2:53	2:56	2:31	1:57	1:54	2:20	n/a	0:17	0:21
Texte et graphismes (5 pages, couleur) Taille du fichier (Ko)	623	924	436	10	19	424	n/a	742	808
Temps de traitement (min:sec)	3:25	3:39	2:53	2:13	2:09	2:19	n/a	0:17	0:21
Feuille de calcul (1 page, noir) Taille du fichier (Ko)	115	59	60	4	4	206	9	25	37
Temps de traitement (min:sec)	0:30	0:39	0:25	0:21	0:23	0:23	0:22	0:07	0:08

⁵ Pour le format HTML, le fichier de sortie est un dossier zippé. Le poids de fichier indiqué est le poids de fichier html réel non compressé avec tous les fichiers graphiques.

Conditions de test des numérisations

Tous les tests de numérisation ont été réalisés sur une HP LaserJet Enterprise flow MFP M525c, avec le micrologiciel FutureSmart version 2 SP1.11 (2201002 231113). Chacun des tests de numérisation figurant dans le tableau n'a été effectué qu'une seule fois et n'est pas le résultat d'une moyenne de plusieurs numérisations. Les informations présentées ne doivent donc pas être considérées comme des données de test exhaustives.

Pour les documents texte et les documents graphiques et texte, nous avons sélectionné trois documents de PAO caractéristiques. Tous ces originaux comportaient cinq pages recto-verso et donc une page blanche en sixième page. La feuille de calcul était sur une seule page, recto uniquement. Les paramètres de numérisation sont présentés ci-dessous.

- Pour toutes les numérisations : numérisation recto-verso, recadrage, suppression des pages blanches, qualité de numérisation moyenne, résolution 150 ppp⁶.
- Pour les documents « Texte uniquement » : noir seul, texte optimisé.

À propos du logiciel HP Digital Sending Software en option

Si les fonctionnalités OCR sont intégrées aux modèles HP LaserJet Enterprise flow MFP et HP Digital Senders, HP propose également HP Digital Sending Software⁷, un logiciel recommandé pour les numérisations vers OCR en volume élevé. HP DSS permet d'optimiser la numérisation d'un grand nombre d'imprimantes multifonctions et de scanners HP (la liste des périphériques compatibles est disponible à l'adresse hp.com/go/dss). HP DSS est un utilitaire d'administration et une interface utilisateur communs à plusieurs appareils et types d'appareils. Ce logiciel en option fonctionne comme un service sur un serveur réseau et permet aux appareils d'envoyer des tâches via le serveur. Il n'est donc pas nécessaire d'installer des logiciels ou des pilotes sur l'ordinateur de chaque utilisateur.

HP DSS présente les fonctions d'envoi indiquées ci-dessous.

- OCR vous permet de convertir des images numérisées dans des formats de fichier courants avec du texte interrogeable et modifiable (les types de fichier sont tous ceux énumérés précédemment dans ce document)
- Copie à distance (envoi à une imprimante)
- Des clés personnalisées permettent aux utilisateurs d'envoyer facilement des documents vers des destinations de flux de travail
- Utilitaire de configuration pour la gestion des envois numériques et des fonctions de flux de travail
- Routage centralisé des e-mails
- Gestion centralisée du carnet d'adresses qui permet aux utilisateurs d'accéder à leurs contacts Microsoft Exchange, ainsi qu'à leur carnet d'adresses privé
- La réplication LDAP permet au DSS de décharger le répertoire d'activité LDAP en répliquant les informations d'adressage utiles du carnet d'adresses DSS
- Prise en charge de Filenet et Novell

Pour de plus amples informations sur HP DSS, consultez hp.com/go/dss.

⁶ Toutes les numérisations OCR sont acquises pour un traitement à 300 ppp, indépendamment des paramètres de résolution en sortie. Mais le fichier qui sera produit sera enregistré selon les paramètres précisés.

⁷ Le logiciel d'envoi numérique HP est une option vendue séparément.

Plus de ressources

Choisir le périphérique HP qui répond à vos besoins en matière de numérisation : <http://h20195.www2.hp.com/v2/GetDocument.aspx?docname=4AA4-3848EEW>

Voir une vidéo de présentation d'une imprimante multifonctions flow : h20621.www2.hp.com/video-gallery/us/en/products/13839A93-A29E-4A02-9EF0-B4A14E0668DC/r/video

Voir une vidéo marketing sur une imprimante multifonctions : youtube.com/watch?v=U0QuJOi0j1Y

En conclusion

La numérisation OCR offre de nombreux avantages, y compris la possibilité de modifier le texte, d'indexer les documents archivés et d'y faire des recherches. L'OCR intégrée aux imprimantes multifonctions HP LaserJet Enterprise flow et à l'appareil d'envoi numérique est encore renforcée par les nouvelles capacités de numérisation de ces appareils multifonctions polyvalents, comme le chargeur automatique de documents de 100 pages, la numérisation recto-verso en un seul passage, la détection de double alimentation par ultra-sons et le traitement avancé de l'image.

Les HP Quick Sets vous permettent de lancer des flux de documents et de le faire bien, en appuyant simplement sur un bouton. Ils automatisent toutes les étapes de ces processus complexes sur un grand écran tactile couleur, dont l'utilisation est un vrai plaisir. Les utilisateurs peuvent trouver immédiatement ce dont ils ont besoin, sans avoir à chercher les réglages appropriés sur le panneau de commande.

Lorsque vous intégrez la numérisation OCR à votre flux de travail, vous devez envisager à la fois les limites de la technologie OCR et les paramètres de performance du matériel. Les informations présentées dans ce document vous aideront à optimiser vos résultats avec l'OCR.

Pour en savoir plus, rendez-vous sur hp.com/go/flow

Plus de ressources

Voir les différentes simulations des imprimantes multifonctions : hp.com/sbso/product/mfp/demo/m575.html

hp.com/sbso/product/mfp/demo/m525.html

Voir une vidéo de présentation d'une imprimante multifonctions flow : h20621.www2.hp.com/video-gallery/us/en/products/13839A93-A29E-4A02-9EF0-B4A14E0668DC/r/video

Voir une vidéo de scanner en réseau d'entreprise : www8.hp.com/h20621/video-gallery/us/en/products/scanners-and-fax/scanners/1377742364001/hp-scanjet-document-capture/video

Voir une vidéo marketing sur une imprimante multifonctions : youtube.com/watch?v=U0QuJOi0j1Y

Abonnez-vous sur hp.com/go/getupdated

© Copyright 2013-2014 Hewlett-Packard Development Company, L.P. Les informations figurant dans ce document sont susceptibles d'être modifiées sans préavis. Les seules garanties pour les produits et services HP sont celles stipulées dans les déclarations formelles de garantie accompagnant ces produits et services. Les informations contenues dans ce document ne constituent en aucun cas une garantie supplémentaire. HP décline toute responsabilité en cas d'erreurs ou d'omissions techniques ou rédactionnelles constatées dans ce document.

Microsoft est une marque déposée du Groupe de sociétés Microsoft aux États-Unis.

4AA4-5159FRE, octobre 2014, rév. 3

