



17/10/06
Dr Rémy L. (Nord)

Optimiser la numérisation de documents

La numérisation des courriers et autres documents externes aux dossiers patients peut représenter plusieurs heures dans la semaine d'un médecin. Comment se simplifier la tâche, quels formats privilégier pour la meilleure intégration dans le logiciel métier ? L'expérience d'un généraliste.

Description

Quotidiennement, je reçois des piles de courriers de mes confrères spécialistes, dactylographiés ou, pire, manuscrits. Comment les incorporer à mon dossier patient informatisé ? Il faut numériser ces documents via un scanner et lire le texte avec un logiciel dédié appelé OCR, pour Optical Character Recognition, c'est-à-dire Reconnaissance Optique de Caractères.

Le texte ainsi récupéré peut alors être placé dans le dossier informatisé, via le module de traitement de texte ou dans un document Word, enregistré à un format universel.

Le matériel



[Cliquez pour agrandir](#)

Les scanners ont beaucoup évolué ces dernières années. On peut simplement prendre un appareil format A4 et numériser chaque feuille : c'est l'opération la plus longue et la plus fastidieuse. Les scanners ont le plus souvent une connexion USB 2.0 et un driver logiciel nommé Twain.

Les médecins qui reçoivent beaucoup de courrier « à scanner » opteront plutôt pour un équipement de type Fuji Scansnap qui présente l'avantage de numériser par lots un ensemble de feuilles. Attention, toutefois, à ne pas y mettre d'agrafes ou de papiers froissés.



[Cliquez pour agrandir](#)

Il existe aussi des stylos scanner, du type IrisPen, très pratiques pour numériser de simples paragraphes à l'instar d'une conclusion ou un compte-rendu opératoire, mais à déconseiller pour des textes longs.



[Cliquez pour agrandir](#)

Les logiciels

Une fois le matériel installé, ainsi que son pilote Twain, vous pouvez commencer à numériser, c'est-à-dire à convertir du texte en images numériques. Ces images peuvent, dans un premier temps, être enregistrées telles quelles, surtout pour du courrier manuscrit.

Le format PDF, pratique et multi plateformes peut être une alternative intéressante. Certains scanners (Canon) offrent un pilote Twain permettant cette conversion directe « IMAGE -> FICHER PDF ».

Sinon, pour enregistrer du texte pur, il vous faudra utiliser un logiciel d'OCR. La plupart du temps, les scanners sont livrés avec un programme « light », épuré de ses fonctions avancées. Néanmoins, il vaut mieux se tourner vers un soft professionnel et rapide. Voici les principaux :

FineReader de Abbyy, société russe, leader mondial des logiciels d'OCR

http://www.france.abbyy.com/finereader_ocr

Readiris de I.R.I.S., société belge, logiciel précis et rapide, OCR en un très grand nombre de langues

<http://www.irislink.com/>

OmniPage de Nuance (ex Scansoft), société américaine

<http://www.nuance.fr/omnipage/>

Intelliant, disponible en shareware

<http://www.intelliant.fr/logiciels-ocr-gestion-electronique-documents.html>

En Opensource ou en freeware :

SimpleOCR (<http://www.simpleocr.com/>)

DocMgr (<http://sourceforge.net/projects/docmgr/>)

GOOCR (<http://jocr.sourceforge.net/>)



Cliquez pour agrandir

Personnellement, j'utilise FineReader 8.0 qui est très rapide et donne une reconnaissance de caractères parfaite, notamment pour les termes médicaux ou pharmaceutiques. Il peut même utiliser le dictionnaire incorporé à Word. On peut également créer des tâches consécutives pour encore plus de gain de temps.

Les formats d'enregistrement

Le format de sortie peut être soit du TXT, mais il ne permet alors pas de mise en page ni de caractères gras, italiques ou soulignés, soit du Word qui facilite la présentation, via un enregistrement sous RTF, DOC ou même HTML.

Si l'on veut conserver l'aspect graphique de la page, on utilisera un format image usuel, comme JPG, GIF ou encore PDF. Éviter absolument les BMP ou TIFF qui ne sont pas compressés et prennent trop de place sur le disque. Un logiciel simple, français et gratuit comme **XNView** permet de lire et enregistrer sous de très nombreux formats, y compris le PDF.

Veiller aussi à utiliser des polices de caractères classiques et bien reconnues par le logiciel de traitement de texte ou par le module interne de texte de votre logiciel métier.

Comment gagner du temps avec l'OCR ?

Reste à optimiser chaque séquence de cette numérisation ou créer des tâches successives automatisées. Le scanner doit être rapide, connecté en USB 2.0 ou encore en SCSI, la définition d'image établie à 300 DPI (vérifier les paramètres du pilote Twain) et le logiciel d'OCR de dernière génération.

Les manipulations sont souvent laborieuses, c'est pourquoi il faut préparer à l'avance sa pile de courriers, retirer les agrafes et trombones, déchiffrer les papiers et numériser d'une seule traite.

Le lot alors enregistré est confié à l'OCR qui va lire intégralement les courriers. Il suffira ensuite d'enregistrer chaque page et de la coller (CTRL+V) dans la page de courrier « reçu » de votre dossier patient en jonglant entre les deux softs ouverts simultanément.

Retour à la liste des astuces

Ce service vous est offert par univadis et les Laboratoires MSD-Chibret. Le contenu de ce service est fourni par DL Infos et ne reflète pas nécessairement l'opinion de univadis ou des Laboratoires MSD-Chibret.
DL Infos 2006