

Les documents numériques

Introduction

Pour certains de nos enfants présentant des troubles spécifiques d'apprentissage, l'outil numérique est incontournable. Pour pouvoir utiliser les logiciels palliant leurs difficultés, il est indispensable qu'ils disposent de documents numériques. Ceux-ci peuvent être créés tels quels ou transformés par numérisation (scanner) et devenir ainsi lisibles, modifiables, stockables sur leur support numérique.

En classe, l'idéal est que les documents de travail soient des documents « PDF ». S'ils sont en papier, ils devront être numérisés sous format PDF. Ceci est possible grâce à un processus complexe appelé « OCR », reconnaissance optique de caractères. Il est à noter que cette étape de numérisation prend du temps, nécessite du matériel supplémentaire si cette fonction n'est pas intégrée dans l'outil (logiciels spécifiques, scanner) et génère une perte non négligeable de la qualité et l'exploitabilité du document.

Dans cette fiche seront présentés ces deux incontournables à l'exploitation numérique que sont le PDF et l'OCR.

1. Les documents PDF

Qu'est-ce que le format PDF (Portable Document Format) ?


Le format PDF est un format de document qui lui permet d'être lu, échangé électroniquement, imprimé, sans que son apparence initiale (mise en page, police de caractère, espace entre les paragraphes, pagination...) en soit changée et ce par **tous** les supports numériques (PC, tablettes, smartphones...). Suivant le programme utilisé il pourra de plus être annoté, copié, lu par des synthèses vocales, transformé en texte...

Les avantages du PDF

- Son universalité : lisible par tous les systèmes (Windows, iOS, Android, Linux.)
- Sa mise en forme « figée » : Le document reste totalement identique au document initial, même après envoi électronique, même si on le complète en dactylographiant contrairement à ce qui se fait dans un traitement de texte classique (si vous ajoutez du texte à un endroit, tout se décale pour « laisser la place » à ce que vous écrivez). Aussi sa pagination reste identique (ce qui est essentiel pour garder des documents numériques de mise en page identique à ceux en papier).

Exemple : Voici une feuille de cours complétée dans Word ou PDF selon le format reçu du document.

Texte fourni par le professeur	texte complété dans Word	texte complété en PDF
Une molécule est constituée d'.....	Une molécule est constituée d'atomes. Les	Une molécule est constituée d' atomes
Les notations chimiques qui rendent compte de la composition des molécules sont appelées.....	notations chimiques qui rendent compte de la composition des molécules sont appelées formules moléculaires.	Les notations chimiques qui rendent compte de la composition des molécules sont appelées formules moléculaires
..... Le chiffre écrit parfois au bas à droite d'un symbole s'appelle	Le chiffre écrit parfois au bas à droite d'un symbole s'appelle indice.	Le chiffre écrit parfois au bas à droite d'un symbole s'appelle indice
Il précise le nombre d'..... de ce type présents dans la	Il précise le nombre d'atomes de ce type présents dans la molécule.	Il précise le nombre d' atomes de ce type présents dans la molécule

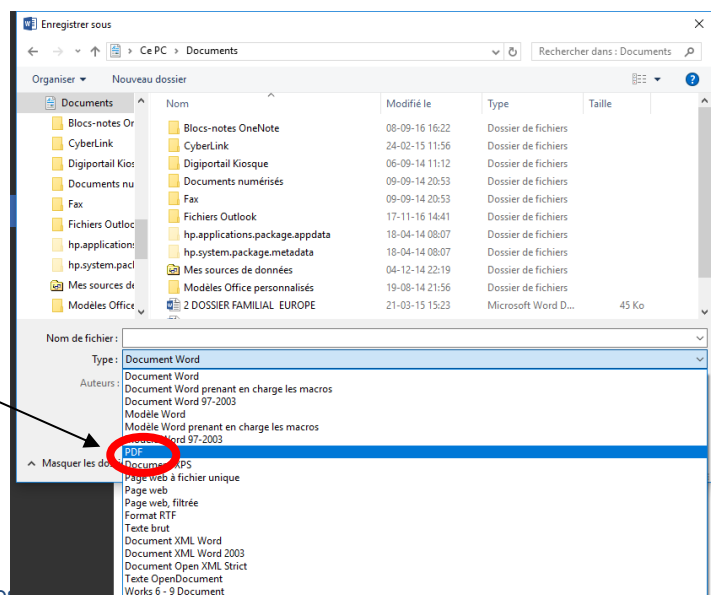
On peut constater que le document complété dans le traitement de texte présente deux inconvénients : sa mise en page est modifiée et les mots rajoutés se confondent dans le texte (si on désire compléter avec une couleur différente à celle du texte initial, il est nécessaire de cliquer sur  à chaque fois que le curseur est mis à un nouvel endroit).

Dans un PDF, la mise en page reste et les annotations se font dans la couleur préalablement prédéfinie aussi longtemps qu'on le désire.

Création d'un document PDF

La plupart du temps les documents PDF sont générés à partir de logiciels de traitement de texte, de tableur, de présentation... Exemple : PDF créé à partir d'un Word :

- cliquer sur « enregistrer sous »
- nommer le fichier dans le dossier de destination
- dans « type », choisir « PDF ».



Sous la coordination d'Anne Floor UFAPEC avenue de
Anne Demanet APEDA – www.apeda.be

Remarque : si la version du traitement de texte est plus ancienne et ne propose pas une telle sauvegarde, il est également possible de créer un PDF en passant par une imprimante virtuelle comme PDF Creator ou en ligne (par exemple sur www.onlineocr.net).

Les documents PDF peuvent aussi être générés à partir d'un scanner auquel le format d'enregistrement à définir est le format PDF.

Il est à noter que si l'outil numérique utilisé par notre enfant ne dispose pas de la fonction scanner intégrée, il devra emmener avec lui en classe un scanner portable dont les plus pratiques et les moins encombrants sont la souris scan (pratique pour des petites surfaces à scanner) ou le scan book.



Remarques :

- Si une adaptation du document est nécessaire (changement de police, une ligne sur deux en couleur, plus d'espace entre les lignes, etc.), elle doit être d'abord pratiquée dans un traitement de texte classique (Word, LibreOffice...). La sauvegarde en PDF se fera ensuite, car de telles modifications sont impossibles sous le format PDF.
- Il est toujours possible de transformer un document réalisé dans un logiciel appartenant à une suite bureautique (traitement de texte, tableur...) en un document PDF. Par contre, transformer un document PDF en un document de type traitement de texte est beaucoup plus complexe, nécessite souvent des logiciels payants et peut entraîner des modifications de mise en page non négligeables.
- Pour des documents de plusieurs pages (volantes), l'utilisation de l'avaleuse de l'imprimante de la maison peut être un gain de temps non négligeable. Seulement le plus souvent elle n'offre pas la possibilité de réaliser du recto verso. Il sera donc nécessaire de trouver un logiciel qui permet de réorganiser les pages. Ainsi *Notability* (sur l'iPad), la version *Pro* payante de *PDF XChange Viewer*¹ le permet facilement. Sinon un logiciel gratuit comme *PDFill FREE PDF Tools* offre parmi ses nombreux outils la possibilité de déplacer des pages au sein des documents PDF.

¹ Voir fiche «Utilisation du logiciel PDF XChange Viewer »

2. L'OCR

L'OCR (Optical Characters Recognition - Reconnaissance Optique de Caractères) est le complément indispensable des documents PDF qui sont issus de documents scannés. La fonction OCR est souvent présente directement dans le scanner mais également dans différents programmes comme PDF XChange ou Dys-Vocal. Cette application signale à l'ordinateur que le document scanné est un texte et non une image. Elle va ainsi chercher à reconnaître chaque lettre avec précision afin de restituer des mots cohérents. C'est pourquoi le programme demande d'indiquer la langue dans lequel le document est écrit.

Utilité

L'OCR est une étape indispensable :

- pour que les textes numérisés en PDF soient exploitables par les différents logiciels d'aide à la lecture et à l'écriture,
- pour pouvoir réaliser des copier-coller (pratique pour différents exercices en classe) au sein d'un PDF,
- pour faire une recherche dans le texte (retrouver un mot spécifique par exemple).

L'OCR rend donc un texte scanné lisible et interprétable par un ordinateur.


Exemples

<p>Voici un article de journal scanné puis traité avec l'OCR :</p> <p>Le Parlement européen et le Conseil de l'Union européenne décident conjointement les actes législatifs qui ont un impact sur le quotidien des citoyens européens. Ces actes portent sur des thèmes comme la libre circulation, la sécurité alimentaire et la protection des consommateurs, l'environnement et la plupart des secteurs de l'économie. Les États membres conservent leur droit de veto dans des domaines tels que la fiscalité et les affaires étrangères/la défense.</p>	<p>Ceci est un « copier-coller » de ce même article dans un traitement de texte :</p> <p><i>Le Parlement européen et le Conseil de l'Union européenne décident conjointement les actes législatifs qui ont un impact sur le quotidien des citoyens européens. Ces actes portent sur des thèmes comme la libre circulation, la sécurité alimentaire et la protection des consommateurs, l'environnement et la plupart des secteurs de l'économie. Les États membres conservent leur droit de veto dans des domaines tels que la fiscalité et les affaires étrangères/la défense.</i></p>
---	---

On peut se rendre compte de ce que l'ordinateur reconnaît comme texte. C'est la partie droite du tableau qui sera lue à haute voix par le logiciel spécialisé.

Certains logiciels spécialisés (payants) permettent la correction de ce « texte sous-jacent ». Un adulte peut donc apporter des corrections pour que le texte soit parfaitement lu. Mais cela demande du temps et de l'énergie et est irréalisable en classe.

Limites

<p>Voici un autre article de journal scanné puis traité avec l'OCR :</p> 	<p>Ceci est un « copier-coller » de ce même article :</p> <p>I e typhon Haiyan qui a frappé les I Philippines les 8 et 9 novembre I derniers fut le plus violent que ce I pays ait jamais connu. Avec des h,ents de plus de 300 krnlh et des vagues de plus de 5 mètres par endroit, il a rasé les îles de Leyte et Samar, au centre de I'archipel, oftant une vision apocalyptique, rappelant les destructions causées par le tsunarni de décembre 2004 en Indonésie. Côté humain,le bilan dépassera sans doute les 5.000 morts. Selon les Nations unies,9 à 13 millions de personnes ont été ûouchées parletyphon ettois mil- lions de pemonnes ont été déplacées.</p>
---	--

Les typographies spéciales, un fond qui n'est pas vraiment blanc, des lignes imprimées de travers, des polices de caractère particulières ou très serrées, ...entraînent inévitablement des erreurs de reconnaissance.

Nous ne le dirons donc jamais assez : l'idéal est d'obtenir un maximum de documents directement par la voie informatique (clé USB, mail, téléchargement) plutôt que de scanner puis corriger fastidieusement les documents scolaires.

Les différentes maisons d'édition francophones l'ont heureusement compris : elles permettent de télécharger gratuitement les manuels scolaires au format PDF à ceux qui en font la demande à condition de fournir une attestation médicale du trouble de l'enfant et, bien sûr, de signer une promesse de non diffusion de ces documents. Ces livres scolaires numériques ne peuvent évidemment servir qu'à l'élève concerné.