

Catalogue des formats de fichiers d'archivage (Cfa) Formats de livres électroniques (E-books) : Analyse de la situation et perspectives

1 Formats de livres électroniques connus

Les formats de livres électroniques se caractérisent par des sauts de page et de ligne dynamiques. Ils ne sont pas axés sur la page et la mise en page comme le PDF ou le XPS. Ils se comportent en principe comme des pages HTML des débuts du web.

A l'heure actuelle, deux formats que l'on peut réellement qualifier de livres électroniques dominent le marché : EPUB et Kindle¹.

- **EPUB** (<http://idpf.org/epub>) est une norme ouverte de l'*International Digital Publishing Forum*.
- **Le format Kindle (KF8)** (https://en.wikipedia.org/wiki/Kindle_File_Format) est un format propriétaire commercialisé conjointement par Amazon et la liseuse Kindle. La lecture est actuellement aussi possible avec des applications.

Tant EPUB que le format Kindle sont disponibles en différentes versions. Il existe des programmes de conversion entre les deux formats.

De plus, les formats suivants jouent également un rôle :

- **DjVu** (<http://www.djvu.org/>) joue un rôle mineur bien qu'il convienne effectivement aussi comme format de livre électronique, en particulier dans le domaine des livres rétronumérisés.
- **iBook**, le format iBook d'Apple est un format EPUB modifié.
- **PDF** avec Adobe-DRM (Digital Rights Management) n'est pas à proprement parler un format de livre numérique, mais il est souvent utilisé sur des liseuses².

Open E-Book et d'autres formats de livres électroniques plus anciens de Microsoft et Sony ne jouent plus aucun rôle à l'heure actuelle.

2 Digital Rights Management (DRM)

Le problème avec les formats de livres électroniques n'est pas tant les formats ou la diversité des formats. EPUB, iBook et Kindle sont basés sur HTML/HTML5 et CSS et peuvent en principe être lus par chaque navigateur. Il est également possible de convertir ces formats en *Plain HTML* sans grande difficulté. PDF et DjVu sont dans ce cas en marge étant donné qu'il ne s'agit pas à proprement parler de formats de livres électroniques et qu'ils ne se basent pas sur du HTML.

Digital Rights Management toutefois pose un problème de principe pour les services d'archives dans la mesure où il limite en règle générale l'utilisation à un utilisateur sur un appareil (ou au maximum deux à trois utilisateurs sur quelques appareils). De plus, DRM peut également limiter la durée d'utilisation. Ce problème ne concerne pas seulement les livres électroniques, mais également des fichiers vidéo et audio. C'est la

¹ Voir <http://www.E-Booknet.de/know-how/E-Book-formate> et https://en.wikipedia.org/wiki/Comparison_of_e-book_formats.

² Voir <http://www.lehmanns.ch/page/E-Bookadobedrm>.

raison pour laquelle les services d'archives n'autorisent généralement par principe aucun fichier DRM protégé.

EPUB, iBook, PDF et Kindle implémentent chacun différemment *Digital Rights Management*, en règle générale en tant que conteneur DRM.

Supprimer la protection DRM est certes techniquement possible, mais n'est pas admissible d'un point de vue juridique. Le service d'archives doit donc s'efforcer à conserver les fichiers sans protection DRM, un problème auquel les bibliothèques sont également confrontées³.

3 EPUB, iBook et Kindle

Si le service d'archives ne trouve pas de voie légale pour traiter des fichiers EPUB, iBook ou Kindle avec DRM protégé, il peut envisager une conversion en HTML ou EPUB étant donné que Kindle⁴ et iBook sont des formats propriétaires et ne sont que partiellement ouverts. Une conversion sans pertes est possible et il existe différents outils pour le faire.

La nécessité de convertir EPUB en HTML dépend de la réduction du nombre de formats dans le service d'archives et de la présence de grandes collections EPUB, par exemple dans les bibliothèques. Une autre conversion d'EPUB/HTML en PDF est également envisageable et possible. Il s'agit dans ce cas d'une véritable conversion et pas uniquement de la suppression d'un conteneur. Une telle conversion a également pour conséquence de perdre l'aspect des sauts de page dynamiques chez les formats de livres numériques.

4 PDF et DjVu

PDF et DjVu sont tous deux des formats axés sur la page et la mise en page. DjVu obtient de bien meilleurs taux de compression pour la rétroumésation de livres⁵, c'est un format ouvert, éprouvé et techniquement plus performant que le PDF. In n'est toutefois par très répandu dans le monde archivistique. C'est la raison pour laquelle il est éventuellement nécessaire de convertir en PDF.

État : 03.04.2017, liens mis à jour en octobre 2021

³ Dans la littérature scientifique, le problème est généralement résolu par le fait que les bibliothèques reçoivent de la maison d'édition scientifique une version avec DRM protégé et la possibilité de recevoir ultérieurement une version non protégée. Une autre possibilité consiste pour la bibliothèque à conclure une convention en ce sens avec l'éditeur du livre électronique garantissant une disponibilité à des conditions restreintes. Voir à ce sujet: <https://www.admin.ch/gov/fr/accueil/documentation/communiques.msg-id-52746.html>

⁴ Le format Kindle a été divulgué par Amazon en 2010 déjà.

⁵ La taille des fichiers DjVu peut être 3 à 5 fois plus petite que le fichier PDF équivalent, voir: <https://en.wikipedia.org/wiki/DjVu>.