

E-books : note sur les formats

Table des matières

E-books : note sur les formats.....	1
<i>Table des matières</i>	1
<i>Un format standardisé : clé du développement des e-books</i>	1
<i>Un format standardisé de digital rights management (DRM)</i>	1
<i>Un forum pour la création d'un standard</i>	2
<i>L'Exemple d'Epub</i>	2
1- Les enjeux.....	2
2 - Les caractéristiques techniques d'Epub.....	2
3 - Le fonctionnement du format EPUB.....	3
4 – Quelques réactions	3
<i>Formats et perspectives</i>	3
<i>Deux sites pour en savoir plus</i>	3

Un format standardisé : clé du développement des e-books

Dans un article d'Avril 2002 "*Standardization aspects of eBook content formats*" publié dans *Computer Standards & Interfaces*, les auteurs, Kyong-Ho Lee, Nicholas Guttenberg, Victor McCrary, identifient la question du format du contenu comme la clé du développement des livres électroniques.

Le livre électronique procure des avantages certains : gain de stockage en place et en coût, communication, livraison et accessibilité multipliée. Il a aussi des inconvénients : qualité d'affichage et de lecture, durabilité et conservation, coût, facilité d'utilisation. Cependant, le manque de format standardisé du contenu reste l'obstacle majeur à son développement.

Le contenu d'un livre électronique peut être constitué dans un format particulier, le format Portable Document Format de l'éditeur Adobe étant alors et encore aujourd'hui le plus largement répandu, et comprendre au sein de ce contenu des liens vers des fichiers multimédia et hypertextuels dans différents formats. Ainsi, un livre électronique peut permettre au lecteur de regarder un clip vidéo, d'écouter du son, ou d'accéder à une page web. Des programmes de traitement vocal du texte ou de traitement en braille élargissent l'accessibilité à des publics malvoyants. Les technologies de la bibliothèque numérique couplées avec les supports mobiles de lecture devraient permettre de multiplier les points d'accès aux contenus des livres et autres documents.

Les différents supports mobiles de lecture offrent différentes fonctionnalités (affichage, navigation, annotation, stockage, etc.). Le format du livre électronique ne peut se contenter d'être conforme à la lecture sur un seul type de support. Il doit pouvoir s'adapter aux différents supports et à leurs fonctionnalités.

Il s'agit d'évoluer de la multiplicité de formats de contenu (PDF, LIT, RTF, HTML, XML, JPEG, TIFF, PDB, PRC, etc.) correspondants à différents types de support¹ vers un format unique standardisé compatible avec tous les supports.

Un format standardisé de digital rights management (DRM)

Les différentes méthodes d'encryptage mises au point par les différents éditeurs rendent difficiles l'avancée vers un standard unique de livre électronique. Sur ce plan là, aussi, la réflexion doit être menée dans le sens de la compatibilité avec les différents supports.

¹ Cf note sur les formats rédigée par la CeB
Le 8 avril 2009 – Cellule e-Book

Un forum pour la création d'un standard

Le format Portable Document Format (PDF) s'est imposé comme un format de lecture de fait. Le logiciel proposé par Adobe, que l'on peut télécharger gratuitement pour une partie des fonctionnalités PDF, n'est pas étranger à ce succès.

Cependant, dès 1998, le National Institute of Standards and Technology (NIST) et les sociétés de logiciels et matériels concernées ont conçu le projet de créer et maintenir un format standard pour le livre électronique.

En 1999, l'Open eBook Forum (OEBF), aujourd'hui International Digital Publication Forum (IDPF <http://www.openebook.org/>), délivre la spécification de l'Open eBook Publishing Structure, aujourd'hui Epub.

Le forum IDPF accueille les éditeurs de livres, revues, journaux, les libraires, les éditeurs de logiciels et de matériels, les auteurs et tous groupes qui ont un intérêt dans les formes de lecture numérique.

L'Exemple d'Epub

Sont présentés ici les principaux enjeux et les grandes lignes des caractéristiques d'EPUB. Pour plus de précisions techniques, se reporter au site de l'IDPF².

1- Les enjeux

Les enjeux que ce format doit relever sont de plusieurs ordres.

a) L'interopérabilité

Le livre électronique doit pouvoir s'échanger entre les divers acteurs indépendamment de leur logiciel et de leur matériel. Epub est donc conçu comme un format non-propriétaire.

b) L'extensibilité

Le livre électronique doit pouvoir évoluer et incorporer de nouvelles fonctionnalités telles des interactions avec le multimédia ou celles avec l'utilisateur. Il faut étudier en particulier la compatibilité entre les générations successives de livres électroniques.

c) L'applicabilité

Le livre électronique doit avoir une interface pouvant s'adapter aux divers systèmes et supports de diffusion, depuis les bases de données aux matériels sans fil à bande passante large comme un serveur internet ou à bande passante étroite comme le téléphone mobile et les assistants personnels. L'intérêt du langage XML -EXtensible Markup Language- a toute son importance ici. Un contenu dans un format XML peut en effet être remodelé, redistribué et grâce à l'utilisation des feuilles de style (CSS) remis en page dans une forme qui convient aux différents supports quel que soit la taille de l'écran et la largeur de la bande passante.

d) L'ouverture

Open source, le standard doit être non crypté, librement accessible à tous les acteurs.

2 - Les caractéristiques techniques d'Epub

Le trait constitutif d'Epub est qu'il se fonde sur les technologies de l'Extensible Markup Language (XML) : XML, XHTML, CSS, MIME, Dublin Core, MARC, Unicode.

Pour répondre aux différentes contraintes, notamment celles d'interopérabilité, d'extensibilité, d'applicabilité, EPUB est composé de 3 standards ouverts :

- Open Publication Structure (OPS)
- Open Packaging Format (OPF)³
- Open Container Format (OCF)

Par ailleurs, Epub offre un mécanisme dit de repli "*fallbacks*" ou de doublage (voir 3).

² <http://www.openebook.org/>

³ Spécifie notamment les métadonnées, les documents constituant l'eBook et l'articulation des documents entre eux.

3 - Le fonctionnement du format EPUB

Le format Epub permet de créer le fichier source du livre électronique et est traduit dans un format natif pour le "reader" avec les DRM propres au pourvoyeur du contenu. Le fichier source peut continuer d'exister ou non dans le fichier qui l'encapsule. Ainsi un fichier source est interprétable par différents supports et plateformes.

Le format Epub permet donc d'être utilisé par tous les lecteurs de livres électroniques.

Il est utilisé pour créer des fichiers source pour les lecteurs suivants :

- Microsoft Reader files : LIT format avec Microsoft DRM
- eBookwise-1150 files : IMP format, IMP DRM
- MobiPocket Reader files : MOBI format : ".mobi" ou ".prc", MOBI DRM
- Amazon Kindle format AZW : MOBI format avec des DRM différents de MOBI
- Peanut Press, Palm Digital Media : .pdb format, eReader.com DRM
- Gemstar (SoftBook press) : .imp format, ETI DRM
- Sony, BBeB/Xylog : .irf format, Sony DRM
- Adobe : .pdf, Adobe DRM
- etc.

En outre, le mécanisme de "fallbacks" d'EPUB exige que pour tout fichier source fourni dans un format propriétaire soit fournie une version de ce fichier dans l'un des 4 formats XML, CSS, JPEG ou PNG. Ainsi, par exemple, un contenu vidéo dans un format propriétaire doit être doublé d'une version JPEG. Cela permet d'assurer l'affichage du contenu par tous les systèmes tout en continuant de permettre que certains contenus ne soient lus que par des systèmes spécifiques. Ce procédé de doublage donne aussi la possibilité d'inclure des versions avec des formats accessibles pour les déficients auditifs ou visuels (cf. Niso et consortium DAISY).

4 – Quelques réactions

La naissance de ce nouveau format a provoqué quelques réactions, permettant de mieux comprendre les attentes et les enjeux :

- Billet **Format ePub sinon rien?** du 26/01/2009 provenant du blog **Aldus – 2006**. Ce billet est un regard critique sur le format EPUB, ses possibilités et ses limites.
http://aldus2006.typepad.fr/mon_weblog/2009/01/format-epub-sinon-rien.html
- Billet du 28/01/2009 provenant du blog de **l'Association Les complexes**.
<http://www.lescomplexes.com/blog/?p=56>
- Billet **Tutoriel de génération de livres au format ePub** du 02/12/2008 du blog **Par-delà**
<http://par-dela.blogspot.com/2008/12/tutoriel-de-gnration-de-livres-au.html>

Formats et perspectives

D'un côté, le format PDF d'Adobe reste le standard le plus utilisé pour les eBooks. Cela tient en partie au fait, déjà signalé plus haut, de la bonne dissémination du logiciel d'Adobe, de son utilisation aisée, de la similarité du fichier PDF avec un exemplaire imprimé. En revanche, l'interaction proposée avec l'utilisateur est limitée à l'annotation. Enfin, le format est réputé pour la protection efficace qu'il apporte contre la reproduction induite.

De l'autre côté, le forum IDPF, auquel Adobe d'ailleurs contribue, paraît très actif. Le format Epub est prometteur. Cependant, on note peu d'éditeurs européens dans la longue liste des membres du forum (<http://www.openebook.org/membership/currentmembers.asp>). Les universités ou institutions universitaires, quant à elles, se comptent sur les doigts d'une main. Une réflexion commune sur ce point, en parallèle à celle de l'offre de contenus, semble pourtant incontournable.

Deux sites pour en savoir plus

Le site des TICE du ministère de l'éducation nationale, de l'enseignement supérieur et de la recherche : <http://www.educnet.education.fr/dossier/livrelec/formats2.htm>

Le site de l'International Digital Publication Forum : <http://www.openebook.org/>