



Assises Professionnelles du Livre A l'heure du numérique

8 novembre 2011 - 14h00 -18h00
Institut océanographique de Paris

Du livre enrichi et de l'EPUB 3

Glossaire des termes techniques utilisés pendant la conférence

ADE (Adobe Digital Editions) : logiciel de lecture permettant de lire des fichiers PDF ou EPUB sur PC ou sur terminal mobile (smartphone, liseuse)

Balises : série de caractères, utilisée pour la structuration d'un élément numérique (texte, images...) et qui sera invisible par le lecteur final. Chaque information est présentée entre deux balises dont le nom donne la nature de l'information. Ainsi, pour déterminer l'élément « Titre », il est nécessaire d'utiliser une balise entrante (<Titre>) et une balise sortante (</Titre>).

Balises <nav> : nouvelle balise introduite dans l'HTML 5 permettant de définir une section dans la navigation (exemples de sections : chapitres, en-têtes, pieds de page,...).

CSS (Cascading Style Sheets) : langage permettant de gérer la présentation d'une page, en donnant des indications comme la police de caractères, la couleur, le positionnement des éléments dans la page, etc. Le but est de séparer la structure d'un document (XML, HTML) de sa forme. Le langage CSS est une recommandation du W3C.

DAISY (Digital Accessible Information System) : le format DAISY est destiné aux déficients visuels ayant des difficultés à lire du texte imprimé. Il est développé par le consortium DAISY, qui en fait la promotion au niveau des standards internationaux. A l'origine, le format DAISY avait pour but d'utiliser les enregistrements audio et d'introduire de la structuration dans les documents afin d'en faciliter la navigation. Depuis, le format DAISY a évolué et offre une expérience de lecture plus flexible et plus agréable pour les handicapés visuels.



Un livre au format DAISY est composé d'un ensemble de fichiers numériques incluant :

- Un ou plusieurs fichiers audio contenant de la narration humaine de tout ou partie du texte d'origine ;
- Un fichier marqué contenant tout ou partie du texte (à strictement parler, ce fichier marqué peut être optionnel) ;
- Un fichier de synchronisation destiné à relier les repères du fichier texte à des repères chronologiques ;
- Un fichier central permettant à l'utilisateur de naviguer entre les différents fichiers tout en maintenant la synchronisation entre le texte et l'audio.

EPUB : A l'origine appelé Open e-book, ce format a pris le nom d'EPUB à partir de sa V2. Il s'agit d'un format de fichier e-book non-propritaire dérivé d'XML, développé par l'IDPF, en voie de devenir le standard pour l'édition électronique. Un fichier EPUB est composé de plusieurs fichiers (instance, structure, métadonnées...), réunis dans un conteneur. En résumé, on peut dire que :

- L'EPUB est un format interopérable de présentation des publications numériques, basé sur la norme XML et les standards du Web.
- L'EPUB permet d'obtenir du contenu web accessible et structuré, interopérable quel que soit le lecteur ou l'application, destiné à un usage en téléchargement ou en streaming.
- L'EPUB est un contenu XHTML comportant une structure de navigation et des métadonnées, le tout intégré dans un package zippé.

EPUB 3 : Il s'agit de la dernière version de la norme EPUB rendue publique par l'IDPF en octobre 2011. Celle-ci intègre des fonctionnalités supplémentaires permettant davantage d'interactivité et de scénarisation. Les principales caractéristiques de l'EPUB 3 sont les suivantes :

- Basé sur l'XHTML 5
- Intègre l'audio et la vidéo (mp3, mp4, aac, atrac)
- Intègre le langage de programmation Javascript
- Support amélioré pour les langues internationales (Japonais, etc.)
- Feuilles de styles et mises en page améliorées via CSS3



- SVG et polices embarquées
- Fonctionnalités améliorées pour les métadonnées
- Supporte MathML (pour les mathématiques)
- Fonction Media overlays (synchronisation audio/vidéo & texte)
- Fonctionnalités améliorées pour la navigation dans le fichier

EPUB Fixed Layout : format proche de l'EPUB 3 et ayant pour caractéristique de figer le contenu, ce qui permet d'afficher l'ouvrage sous forme de double page comme le livre papier. Il est en voie de normalisation par l'IDPF, mais est déjà supporté par certaines tablettes, dans des versions qui leur sont spécifiques (exemple : iPad).

HTML : c'est le format de données utilisé pour représenter les pages web. Il se présente sous la forme d'un ensemble de balises permettant de structurer et de mettre en forme le contenu des pages et d'y inclure des ressources multimédias comme des images ou des formulaires de saisie par exemple.

HTML5 : ensemble de balises qui permettent de décrire plus précisément le contenu des pages web, il introduit de nouvelles notions comme l'article, la section, l'en-tête ou le pied de page. Cette norme, encore en cours d'évolution, devrait être finalisée vers 2014.

iBooks : logiciel de lecture d'Apple disponible sur iPad et iPhone permettant de lire des livres numériques aux fichiers PDF et EPUB notamment.

IDPF (*International Digital Publishing Forum*) : Cet organisme international a pour vocation de favoriser au niveau global la disponibilité, l'enrichissement et l'accessibilité des publications numériques en développant et en promouvant des standards ouverts et interopérables. Il compte parmi ses adhérents des grands noms de l'Internet, des constructeurs de matériels, des marques de logiciels, des éditeurs, des revendeurs,...



Javascript : langage de programmation de scripts incorporé dans un contenu HTML, permettant de développer des services, des applications ou des programmes à destination du Web notamment.

MathML : définie par le W3C, cette spécification permet de définir le langage mathématique en intégrant des fonctionnalités de tableaux et de diagrammes.

Métadonnées : Terme générique pour toute information descriptive du produit concerné. Elles sont un élément essentiel de l'architecture Web. Elles sont essentielles pour toute communication sur le produit livre quel que soit le support (un catalogue papier n'est fait que de métadonnées). Elles permettent en particulier d'identifier et de décrire les ressources documentaires, par exemple :

- contenu : titre, sujet, description, source, langue, relation, couverture ;
- propriété intellectuelle : créateur, éditeur, contributeur, droits (droits d'auteur, etc.) ;
- matérialisation : date, type, format, identifiant.

Mobipocket : A l'origine, nom de la société ayant mis au point le format « .mobi ». Ce format de fichiers correspondant à des livres numériques est spécifique à Amazon depuis le rachat de la société Mobipocket par la firme de Seattle. Ce format devrait être remplacé à terme par le nouveau format Kindle Format 8 (KF8) annoncé par Amazon en octobre 2011.

Mode vectoriel : technologie qui permet d'agrandir une image sans qu'elle se pixellise.

PDF (*Portable Document Format* – format de documents transférables) : Langage de description de pages pour l'impression créé par Adobe. Il présente l'avantage de préserver les polices, les images, les objets graphiques et la mise en forme, quelles que soient l'application qui a servi à le créer et la plateforme utilisée pour le lire. www.adobe.com



SVG : Ce format de fichiers images basé sur le langage XML et recommandé par le W3C figure dans les spécifications de l'IDPF pour le format EPUB 3. Il permet notamment d'introduire des images sous forme vectorielle sans pixellisation ou de les redimensionner sans perte de qualité.

World Wide Web Consortium (W3C) : organisme international créé en 1994, qui développe des standards pour le Web afin de faciliter la communication par Internet, indépendamment du matériel, du logiciel, de la langue ou du handicap. www.w3.org.

XML (*eXtensible Markup Language* – langage de balisage extensible) : Le XML est un langage standard de balisage générique : chaque information est présentée entre deux balises dont le nom peut être choisi. Il favorise l'échange automatisé d'informations et de contenus entre des systèmes d'informations hétérogènes (interopérabilité). Le jeu de caractères du langage XML est Unicode. C'est une recommandation du W3C, <http://www.w3.org/TR/REC-xml>