

Producción de  
contenidos  
estructurados.  
Necesidades  
metodológicas.

## XML y el libro electrónico

Ricardo Eito Brun  
Madrid, 1 de junio 2015

---

# El libro-e.

## Notas sobre su evolución

- La edición de libros en formato electrónico es una realidad consolidada, tanto en el ámbito académico como en el “mercado general de gran consumo”.
- En un primer momento, papel predominante la edición científico-técnica (Springer Link, InterScience, etc.)
- Actualmente, popularización de dispositivos lectores tanto especializados (Amazon Kindle) como genéricos (Ipad).
- Incorporación de libros-e en bibliotecas (públicas y universitarias), no sólo a través de las “grandes bases de datos”.
- Plataformas de gestión de libros-e (XeBooks, Libranda, etc.) para bibliotecas públicas.
- Autoedición.

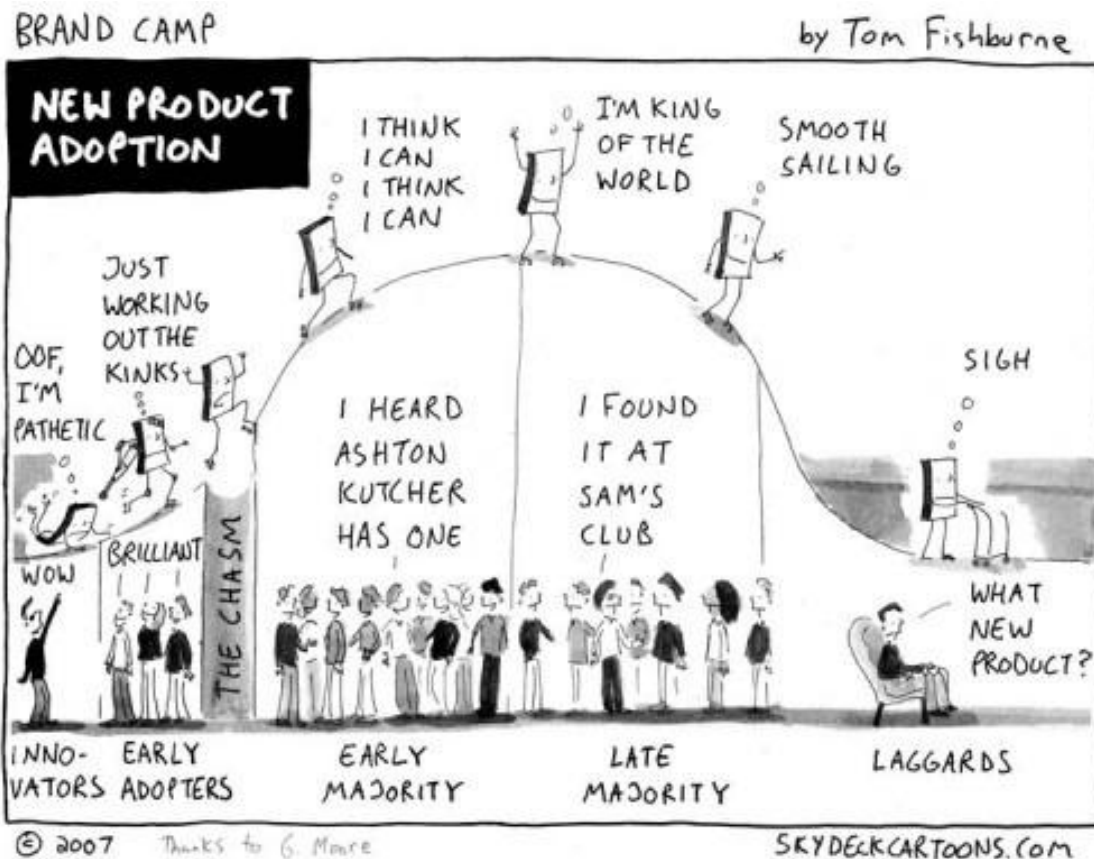
# El libro-e.

## Notas sobre su evolución

- El modelo se ha definido tanto desde su perspectiva técnica como comercial o de mercado.
- Evolución y muerte (Microsoft LIT, Glassbook, Gemstar, Franklin Publishing, etc.) de algunas iniciativas tecnológicas:
- Tras la consolidación técnica, se produce la adopción generalizada.
- ¿En qué punto estamos?

# El libro-e.

## Notas sobre su evolución



## El libro-e. Hitos (1ª etapa)

- 1998 – Comercialización dispositivos dedicados. Rocket ebook y Softbook.
- Agosto 2000 - Adobe compra Glassbook, creada en 1999 para ofrecer a editoriales software de distribución / comercialización segura para libros-e
- 2000 - Acuerdos estratégicos con barnesandNoble.com, Amazon, WaterStone, etc
- Mayo 2002 – Planeta cierra veintinueve.com
- Marzo 2003 – Gemstar anuncia el final de sus dispositivos de lectura (basados en Rocket eBook de Nuvomedia y SoftBook Press).
- Agosto 2011 - Microsoft anuncia retirada de MS Reader.

## El libro-e. Hitos (2ª etapa)

- 2005 – Amazon compra “MobiPocket”.
- 2006/07 - Sony, Amazon y Papyre lanzan al mercado sus dispositivos basados en tinta electrónica, sin retroiluminación.
- 2006/10 - Teléfonos móviles equipados con pantallas de gran formato y lectores de libros-e.
- 2008 – Sony avanza hacia la compatibilidad (BBeB > Epub).
- 2010 - tablet táctil de Apple y NookColor de B&N.
- 2010 - Association of American Publishers, el negocio en USA alcanzó 316,6 Meuros, 8,3% de la facturación total.
- 2010 - En España se comienzan a popularizar los dispositivos, sin una disponibilidad de contenidos equivalente.
- 2013 – Google Play Books distribuy libros-e.
- Primeras plataformas: Publidisa (Todoebook), 36L Books, Amabook, Leer-e y Libranda.

## El libro-e. Notas sobre su evolución

- Bibliotecas Públicas Estado: “Proyecto de adopción, con una primera fase (2010/11) para promover y difundir el uso del libro-e en las BPE.”
- Presupuesto : 126.824 €
- Compra de 677 dispositivos
- Se distribuyeron en 15 bibliotecas: Ciudad Real, Huelva, La Rioja, Las Palmas de Gran Canaria, Madrid, Mérida, Murcia, Oviedo, Palma de Mallorca, Santander, Santiago de Compostela, Tarragona, Valencia, Valladolid y Zaragoza,
- Cada biblioteca recibió una media de 45 e-readers.
- Los dispositivos seleccionados fueron :
  - Papyre de Grammata (600 títulos)
  - Cooler de Cool-ER (500 títulos)
  - Wolder MiniBuk 5”, de Wolder 1.000 títulos
  - Cybook Opus, de Booken, 100 títulos
  - Book Cervantes, Luarna 1.000 títulos
  - Inves Wibook 600, 184 títulos

# El libro-e.

## Notas sobre su evolución

- Estos e-readers venían precargados con obras de dominio público
- El número total de títulos disponibles ha sido de 3.384 títulos.
- La experiencia de estas bibliotecas permitió al Ministerio de Cultura evaluar el servicio y la acogida de la iniciativa por parte de los usuarios durante los tres primeros meses de préstamo, de enero a marzo de 2011.
- Los datos se extrajeron de 439 encuestas recibidas en estos primeros meses.



# El libro-e.

## Notas sobre su evolución

- La valoración del servicio:
  - El 36% de los usuarios considera que es muy bueno,
  - el 28% excelente,
  - el 19% bueno,
  - y el 17% regular,
  - ningún usuario manifiesta que el servicio sea malo.
- ¿Los títulos disponibles en el dispositivo han resultado interesantes?
  - Suficientemente interesantes e interesantes, a un 31% y 30%
  - Poco interesantes, 27%
  - No interesantes, 2%.

# El libro-e.

## Notas sobre su evolución

- Dificultad encontrada en el manejo de este tipo de herramienta,
  - 41% de los usuarios manifestaron no haber tenido dificultades y que era manejable,
  - 26% muy manejables,
  - 15% suficientemente manejable,
  - 11% dificultosa,
  - 7% extremadamente dificultosa.
- La segunda fase estaba prevista para 2012, con una inversión de 165.870 €, para adquirir 1.391 dispositivos con 6.284 títulos.
- Los modelos de e-readers a evaluar eran: Booq Cervantes Lite, Papyre 6.1., iCarus Go, iCarus Sense, Inves Wibook 650T, Wolder Mibook Sensations, Tagus, Booq Avant 3 y Booq Cervantes 2.

# El libro-e.

## Notas sobre su evolución

- Accesibilidad que permita el acceso en línea desde cualquier lugar.
- Calidad de la plataforma, que permita una gestión de los recursos ágil e integrada con el OPAC de la biblioteca.
- Facilidad de uso de la interfaz y que incorpore prestaciones 2.0.
- Posibilidad de trabajar con diferentes tipos de formatos de lectura (PDF, EPUB, MOBI).
- Disponibilidad del préstamo y de la lectura de los contenidos en cualquier dispositivo (ordenadores personales, lectores de libros electrónicos, tabletas, teléfonos inteligentes, etc.), ya sea a través de descargas o en tiempo real.

# El libro-e.

## Notas sobre su evolución

- En las bibliotecas norteamericanas se evidencian posturas contrarias al papel de OverDrive.
- La red Douglas County Libraries, Colorado, ha apostado por un modelo en el que, desarrollando su propio software, puedan comprar libros directamente de las editoriales, conseguir mejores precios y condiciones y garantizar herramientas de búsqueda que permitan leer en línea los libros.
- Este mismo esquema lo siguen también las bibliotecas de California dentro del consorcio Califa Library Group, que se plantea comprar archivos de libros electrónicos directamente a los editores ya que su objetivo es tener la propiedad sobre los libros electrónicos.

## El libro-e. Notas sobre su evolución

- Bibliotecas Públicas Estado: eBiblio
- Presupuesto : 1.604.700 €
- Iniciado en septiembre 2014
- Se ofrece acceso a bibliotecas públicas de distintas comunidades a una “plataforma de préstamo de libros-e”
- En diciembre 2014, se alcanzaron 42.216 préstamos de un total de 1.500 títulos.
- Los resultados se consideran positivos.
- Inicialmente se trabaja con Librandia, pero cada CC.AA podrá elegir otros proveedores adicionales.

# El libro-e.

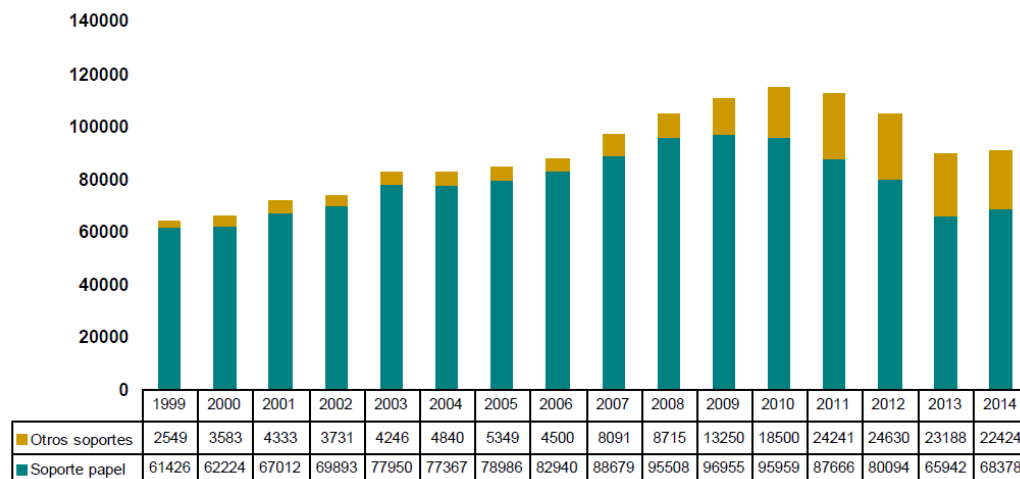
## Notas sobre su evolución

- ¿Las bibliotecas como impulsoras y creadoras de contenido desde una doble perspectiva?
- Estimular a los diferentes servicios de publicaciones de cada administración y participar en los proyectos de plataformas de préstamo de libros electrónicos.
- Ello serviría para contribuir a difundir este tipo de publicaciones, que ganarían en visibilidad, y a la vez aumentaría la oferta de la propia biblioteca, con temas y contenidos que de otro modo pasarían inadvertidos para un gran número de usuarios.
- la biblioteca, dentro de su capacidad de formación en nuevas tecnologías, podría colaborar para impulsar la creación de contenidos locales.
- Estos contenidos pueden no ser atractivos para la industria editorial .

# El libro-e.

## Notas sobre su evolución

- “La edición de libros digitales muestra respecto al año anterior un descenso del 1,9%, con 20.263 ISBN. Su peso sobre el total de la oferta retrocede del 23,2% al 22,3%. Se invierte por primera vez la tendencia iniciada en 2008 hacia el aumento de la edición digital en detrimento de la impresa” \*



\*El sector del libro en España (2013-2015) del Observatorio de la Lectura y el Libro (MECD)

# Libros-e: formatos

- Al principio, las principales opciones eran PDF y OEB (Open eBook)
- **PDF:**
  - Formato de “réplica”
  - No es re-paginable, impide que se puedan adaptar correctamente a la pantalla del dispositivo lector.
  - ISO 32000-1:2008 y “familia” de estándares PDF/A.
  - Más complejo de procesar.
- **OEB**
  - Especificación abierta para garantizar la creación de libros-e independientes de plataformas específicas
  - Primera versión 1.0 publicada en septiembre 1999
  - Participaron: Adobe, Andersen Consulting, Microsoft, Palm, Nuvomedia, IBM, IDB, McGrawHill, Time Warner, EveyBook, Barnes & Noble, Softbook Press, etc.



# Libros-e: formatos. OEB

- Basada en HTML, XML, Dublin Core
- Soportaba imágenes JPEG y PNG, y hojas de estilo CSS.
- Un software de lectura podía ampliar el conjunto de etiquetas “básicas” del OEB con formatos y etiquetas adicionales, y seguir siendo “OEB-compliant”
- Elegida por Microsoft (.lit era OEB compilado) y por los “dispositivos dedicados).

# Libros-e: formatos. OEB

## OEB y XML

- Un libro-e OEB contaba siempre con un archivo en el cual se hace referencia a:
  - Los distintos archivos HTML que formaban el texto del libro-e final
  - El orden en el que se presentaban estos archivos
  - Secuencias de lectura adicionales
  - Imágenes y otros archivos que se utilizaban en el libro-e
  - Metadatos descriptivos e identificativos.
- Esta información se guardaba en un archivo XML con extensión .opf – Open eBook Package File -.
- El archivo .opf era un archivo XML válido, basado en la DTD oebpkg1.dtd

# Libros-e: formatos. OEB

## OEB y XML

- El elemento raíz era `package`
- `package` tenía un atributo `unique-identifier`, que tomaba como valor un ID único para el libro-e
- El elemento `package` contenía:
  - Un elemento `metadata`
  - Un elemento `manifest`
  - Un elemento `spine`
  - Un elemento `guide`

```
<package unique-identifier="84-567-45-23">  
    <metadata> ..... </metadata>  
    <manifest> ..... </manifest>  
    <spine> ..... </spine>  
    <guide> ..... </guide>  
</package>
```

# Libros-e: formatos. OEB

## OEB y XML

- Contenía metadatos del libro-e: título, autor, materia, editor, fecha publicación, etc.
- Incluía metadatos Dublin Core, dentro de un elemento `metadata`, sólo dos de ellos obligatorios:
  - Elemento `dc:identifier`
  - Elemento `dc:title`
- Además de los metadatos del DC, se podían utilizar metadatos adicionales en un elemento `x-metadata`

# Libros-e: formatos. OEB

```
<package unique-identifier="84-5455-111-32">
<metadata xmlns:dc="http://purl.org/dc/elements/1.0/"
xmlns:oebpackage="http://openebook.org/namespaces/oeb-
package/1.0/">
  <dc-metadata>
    <dc:Title>Viaje a la Alcarria</dc:Title>
    <dc:Identifier>84-5455-111-32</dc:Identifier>
  </dc-metadata>
  <x-metadata>
    <meta name="signatura" content="34CELcam">
    <meta name="acceso" content="publico">
  </x-metadata>
</metadata>
.....
</package>
```

## Libros-e: formatos. OEB

OEB y XML: manifiesto

- Contenía la lista de archivos (texto, imagen, etc.) que conformaban el libro-e.
- Se podían indicar en cualquier orden.
- Cada archivo se indicaba en un elemento vacío ítem.
- Ítem aceptaba tres atributos:
  - Id – identificador único para el archivo. Se le hará referencia en otras secciones del archivo .opf
  - href – ruta física del archivo, completa o relativa a la situación del archivo .opf
  - media-type – tipo MIME del archivo, por ejemplo text/x-oeb1-document, image/png, contents/xml, etc.

# Libros-e: formatos. OEB

```
package unique-identifier="84-5455-111-32">
<metadata>
.....
</metadata>
</manifest>
  <item id="cap1" href="capitulo_1.htm" type="text/x-oeb1-
document" />
  <item id="cap2" href="capitulo_2.htm" type="text/x-oeb1-
document" />
  <item id="cap3" href="capitulo_3.htm" type="text/x-oeb1-
document" />
  <item id="cap1i11" href="img/img1.png" type="image/png" />
  <item id="cap1i12" href="img/img2.png" type="image/png" />
  <item id="cap2i11" href="img/img3.png" type="image/png" />
</manifest>
</package>
```

# Libros-e: formatos. OEB

## OEB y XML. SPINE

- Determinaba el orden de lectura principal de la publicación, es decir, en qué orden debían aparecer los archivos indicados en el elemento manifest.
- El elemento spine contenía uno o más elementos itemref.
- El elemento itemref es un elemento vacío, que tenía un atributo idref que tomará como valor el del atributo id de los elementos item de la sección manifest al cual hacían referencia.

```
<spine>  
  <itemref idref="abc" />  
  <itemref idref="abd" />  
</spine>
```



# Libros-e: formatos. OEB

## OEB y XML. Guide

- Elemento opcional, establece “componentes estructurales” del libro-e: cubierta, glosario, sumario, índice, lista de tablas, lista de imágenes, etc.
- El elemento guide contendrá un elemento reference para cada componente estructural.
- El elemento reference es un elemento vacío, que tiene tres atributos:
  - type – tipo de elemento
  - title – Nombre descriptivo para el elemento, tal y como lo verá el lector
  - href – URL absoluta o relativa del archivo correspondiente al elemento estructural. Debe ser un archivo declarado en la sección manifest.

## Libros-e: formatos. OEB

- El atributo type del elemento reference puede tomar uno de los siguientes valores:
  - cover – cubierta del libro
  - title-page – página con el título, autor, editor, etc. (portada)
  - toc – tabla de contenidos o sumario
  - Index – índice del libro
  - glossary – glosario
  - acknowledgements – página de agradecimientos
  - bibliography – bibliografía
  - colophon – colofón
  - copyright-page – página con noticia del copyright
  - dedication – dedicatoria
  - loi y lot (lista de imágenes y de tablas)
  - epigraph, foreword, preface y notes

# Libros-e: formatos. OEB

```
<package unique-identifier="84-543-5454-2">
.....
<guide>
  <reference type="toc" title="Sumario"
  href="toc.htm" />
  <reference type="loi" title="Im&225;genes"
  href="images.htm" />
  <reference type="lot" title="Tablas"
  href="tablas.htm" />
</guide>
</package>
```

Remplazado por `<nav type="landmark">`

## Libros-e: formatos. OEB

- La especificación OEB se basaba en HTML.
- Los contenidos se debían crear o convertir a este formato
- Cada documento HTML debe ser un **documento XML bien formado**:
  - En los elementos vacíos se debe usar la sintaxis `<... />`  
`<img href="hola.gif" />`
  - Los elementos deben estar anidados correctamente
  - El documento debe comenzar con una declaración XML  
`<?xml version="1.0"?>`
  - Los elementos no vacíos deben contar con una etiqueta de inicio y otra de fin.
  - El valor de los atributos se debe escribir entre comillas dobles.

## Libros-e: formatos. OEB (diferencias html)

- Existían diferencias entre el “HTML convencional” y la versión de HTML que utiliza OEB
  - Los elementos utilizados para “aplicar formato visual” se eliminaron o su uso de desaconsejó a favor de hojas de estilo CSS
  - Los atributos id, style, class y title eran aplicables a todos los elementos

## Libros-e: formatos. OEB (diferencias html)

- `<a>...</a>`
  - origen y destino de enlaces
  - Atributos href, rel (name se elimina)
- `<body>...</body>`
  - Cuerpo del documento
  - Atributos bgcolor, text, vlink, link, alink se eliminan
- `<caption>...</caption>`
  - Título de tablas, el atributo align se elimina
- `<center>...</center>` .... Se elimina
- `<img />`
  - Inclusión de imágenes
  - Los atributos width, height, align, border, hspace... se eliminan

## Libros-e: formatos. OEB (diferencias html)

- `<ol>...</ol>`
  - Listas ordenadas o numeradas
  - Atributos `type`, `start` se eliminan
- `<p>...</p>`
  - Párrafos
  - Atributo `align` se elimina
- `<s>...</s>`, `<u>...</u>` - se eliminan
  - Título de tablas, el atributo `align` se elimina
- `<style>...</style>`
  - El atributo `type` debe recoger el valor: `text/x-oeb1-css`
- `<table>...</table>`
  - Inclusión de tablas
  - Los atributos `align`, `bgcolor` se eliminan

## Libros-e: formatos. OEB (CSS)

- La especificación OEB utilizaba un modelo “simplificado” de hojas de estilo CSS
- Selectores posibles:
  - ELEMENTO { }
  - ELEMENTO.valorClass { }
  - .valorClass { }
- Los pseudo-selectores, los basados en ID o en contexto no son exigidos por la especificación
- Es posible agrupar selectores en una misma regla, separándolos mediante comas:
  - H1, h2, h3 { font-family : Times }



## Libros-e: formatos. OEB (CSS)

- Para aplicar estilo con CSS a un documento HTML, la especificación OEB permitía utilizar :
  - Hojas de estilo externas, unidas mediante el elemento LINK
  - Hojas de estilo internas, definidas con el elemento STYLE
  - Utilización del atributo STYLE
- Otros cambios:
  - Propiedad display, se añaden valores oeb-page-head y oeb-page-foot
  - Propiedad font-weight, acepta sólo los valores normal y bold (no acepta escala 100 a 900 ni valores relativos bolder y lighter)
  - La propiedad list-style-type – no acepta los valores disc, circle y square
  - Propiedad text-decoration no acepta el valor line-through

## Libros-e: formatos. EPUB

- Se considera “heredero” y evolución de OEB.
- International Digital Publishing Forum gestiona su evolución.
- EPUB2 en 2007.
- EPUB2 basada en XHTML 1.1, CSS2 y XML para definir la estructura de la publicación.
- En 2008, la AAP declaró su apoyo a este formato,
- Respaldado por editoriales como Penguin, HarperCollins, Cambridge University Press, John Wiley & Sons, Hachette, Oxford University Press, Random House, etc.
- Por parte de la industria:
  - Sony adoptó este formato en sus lectores.
  - Adobe incluyó un proceso para generar EPUB desde Indesign.
- EPUB3 en 2011.
- En 2013 la AAP lo convirtió en estándar y recomendó a sus afiliados digitalizar sus obras en EPUB3.

## Libros-e: formatos. EPUB

¿Qué diferencia EPUB3 de otros formatos?

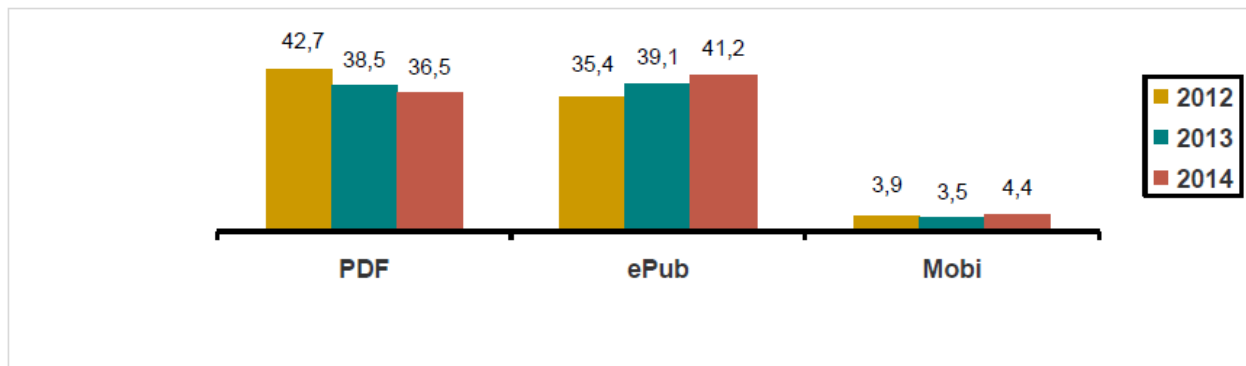
- Accesibilidad:
  - Funcionalidad (text-to-speech)
  - Conversión de vídeo a texto.
  - Soporte a las lenguas no occidentales
- Adaptación a dispositivos:
  - Se adapta a distintos tamaños de letra y pantalla.
  - Se puede cambiar el tipo de letra.
  - El formato marca el contenido, pero no delimita su formato.
  - “Fixed-layout” para libros infantiles.
- Independencia

## Libros-e: formatos. EPUB

- La especificación no depende de ninguna empresa particular.
- Se reducen costes para el usuario final.
- Más actores pueden desarrollar aplicaciones y utilidades software (conversión, visualización, etc.)
- Estrategia opuesta a la de Amazon con sus formatos para Kindle.
- EPUB ha sido “adoptado” por empresas creadoras de dispositivos lectores como Sony, Apple y soportado por Adobe.
- Adobe DRM es el más utilizado (junto con Apple)
- DAISY Consortium adoptó en 2010 EPUB3 como alternativa para audiolibros frente a NISO Z39.86.

## EPUB. Presencia e importancia

- “En cuanto a la oferta por formatos digitales, según el avance de **Panorámica de la Edición Española de Libros, en 2014 crece por segundo año consecutivo la oferta en formato ePub, superando los libros publicados en PDF, que representan el 41,2% y el 36,5% respectivamente. Esta tendencia también se reflejó en los datos de 2013 del mismo informe**”



Fuente: Panorámica de la Edición Española de Libros. Avance 2014

\*\* El sector del libro en España (2013-2015) del Observatorio de la Lectura y el Libro (MECD)

## EPUB. Presencia e importancia

- “La aplicación de un formato u otro varía en función de las materias. Así, **en el caso de los libros de Infantil y Juvenil, Creación Literaria, Tiempo libre y Otros**, el formato **ePub se aplica con mayor frecuencia** que el PDF. Por el contrario, en los libros de Ciencias Sociales y Humanidades, Ciencia y Tecnología, y Texto, se está utilizando más el formato PDF.”\*\*

Tabla 45 – Clasificación de libros digitales según subsectores y formatos, 2014

	PDF	ePub	Mobi	Otros	Total
Infantil y juvenil	134	856	90	36	1.116
Texto	1.003	80	2	2.949	4.034
Creación Literaria	1.095	3668	304	80	5.147
CC Sociales y Humanidades	3.569	2267	323	291	6.450
CC y Tecnología	1.087	721	40	214	2.062
Tiempo Libre	385	543	88	32	1.048
Otros	118	217	48	23	406
<b>Total</b>	<b>7.391</b>	<b>8.352</b>	<b>895</b>	<b>3.265</b>	<b>20.063</b>

Fuente: Panorámica de la Edición Española de Libros. Avance 2014

\*\* *El sector del libro en España (2013-2015) del Observatorio de la Lectura y el Libro (MECD)*

## EPUB. Presencia e importancia\*

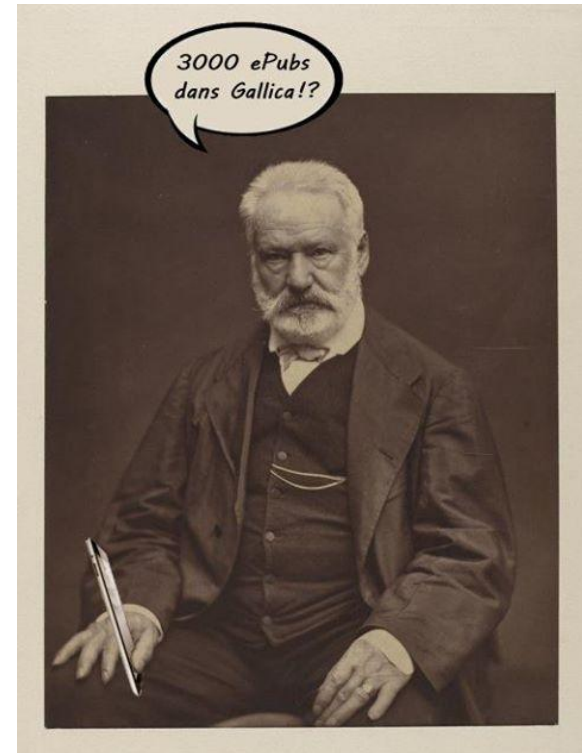
- **4. El triunfo del formato ePub**
- Durante los años 2009 y 2010 comprobamos cómo la oferta de libros electrónicos en nuestro país estaba constituida, principalmente, por libros en formato PDF.
- Las últimas cifras publicadas por la Agencia ISBN relativas a 2011 ponen en evidencia que a lo largo del pasado año se realizó una apuesta generalizada por ePub, descendiendo la aplicación del PDF.
- Durante ese periodo 6 de cada 10 ebooks empleaban ePub.
- La adaptación automática del texto en ePub al tamaño de la pantalla facilita su lectura en casi todos los dispositivos —excepto Kindle—.

\* Situación actual y perspectivas del libro digital en España II. La producción española de libros digitales y su distribución y venta en la Red. Observatorio de la Lectura y el Libro (MARZO 2012)

## EPUB. Presencia e importancia BnF (Gallica)

*« Le format ePub, compatible avec la grande majorité des tablettes et liseuses actuellement sur le marché, offre l'avantage d'être un format léger qui prend moins de place qu'un PDF.*

*Il permet également au texte de s'adapter à la taille des écrans, ce qui favorise la lecture et l'accessibilité au texte pour les non-voyants et mal-voyants, comme on peut le constater... »*





# EPUB. Presencia e importancia Depósito legal (BL.UK)

*“For electronic publications published in more than one digital format, it may be necessary to agree which digital format should be deposited. Common accepted formats include XML, HTML, SGML, PDF, EPub, Microsoft Word and RTF.*

*What options are available for delivering content to the Legal Deposit Libraries?*

*For publishers of books, journals or other documents in PDF, Microsoft Word, RTF, EPub and other non-XML formats...”*

<http://www.bl.uk/aboutus/legaldeposit/websites/elecpubs/>

The logo for the British Library, featuring the words "BRITISH LIBRARY" in white, uppercase, sans-serif font, stacked vertically on a red rectangular background.

BRITISH  
LIBRARY

## EPUB. Presencia e importancia bne.es

*“Todo lo que se publica en línea es objeto de depósito legal, independientemente del formato. Sea cual sea el formato, es obligación de los centros de conservación garantizar la preservación de las publicaciones y el acceso a ellas.”*



Distintos libros disponibles en formato epub.

[www.bne.es/webdocs/Servicios/.../DLe\\_PreguntasFrecuentes.pdf](http://www.bne.es/webdocs/Servicios/.../DLe_PreguntasFrecuentes.pdf)

# EPUB. El formato

- Especificación 3.0.1, 26/06/2014.
- Un libro-e EPUB es un fichero zip que cumple con 3 especificaciones:
  - OCF. Open Container Format. (formato archivo para encapsular un conjunto de recursos en un zip)
  - Publications (Open Packaging Format ).
  - Content Documents (Open Publication Structure).
- En el nivel de raíz debe existir un directorio META-INF con un archivo container.xml.
- Puede incluir firmas y datos de encriptación.
- Contenido HTML5 (serialización XHTML)
- Media Overlays – para sincronización texto-audio.
- Core Media Types: gif, peg, png, svg, xhtml, mp3, fuentes openType y WOFF.

# EPUB. El formato

- OPF (Open Packaging Format).
  - Archivo de metadatos XML.
  - Especifica la ubicación de todo el contenido del libro y los recursos: textos imágenes, etc.
  - Incluye metadatos descriptivos Dublin Core, con 4 elementos obligatorios.
  - El manifiesto lista todos los recursos que forman el libro-e (lista de archivos XHTML, imágenes, CSS, etc.). Todos los archivos que conforman el libro-e deben listarse en el manifiesto.
  - “Spine”. Indica el orden de lectura lineal del libro-e.
  - Tabla de contenidos.
  - Fallbacks y bindings.

# Otros formatos. Amazon

- Amazon ha desarrollado formatos propietarios para su lector Kindle.
- KF8 o AZW3 (Kindle Format 8), publicado a finales de 2011.
- También basado en HTML5 y CSS3, con una estructura similar a EPUB.
- Características adecuadas a libros-e para niños (fixed-layout)
- Basados en el formato mobi de Mobipocket.
- Fácil conversión desde EPUB a KF8 con KindleGen.



## Amazon Kindle Publishing Guidelines

How to make books available for the Kindle platform

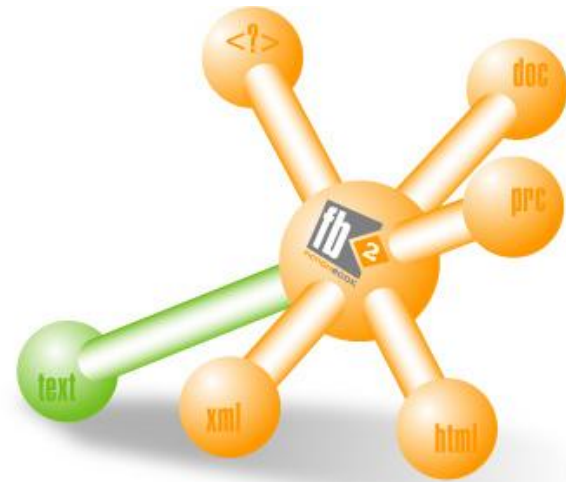
version 2015.1

This document describes the primary ways publishers, authors, and conversion houses can make their content available on the Amazon Kindle platform. This document includes guidelines and suggestions to ensure a smooth conversion and publication process.

Copyright © 2015 Amazon.com, Inc. or its affiliates. All rights reserved.

## Otros formatos. fb2

- FictionBook basado en XML para el almacenamiento de libros donde cada elemento del libro es descrito por etiquetas.
- Un libro-e consiste en un único archivo XML donde se incorporan las imágenes.
- Se reconoce “apto” para libros de ficción, pero limitado para contenidos científico-técnicos.
- Ampliamente usado en Rusia (Donde se creó)



# Otros formatos. DJVU y CBA

- DJVU
  - Formato libre centrado en imágenes escaneadas con resolución “alta” (300-400dpi).
  - Sistema optimizado de compresión desarrollado por AT&T Labs en 1996.
- Comic Book Archive
  - Ideado para “comics”.
  - Conjunto de imágenes (png, jpg) ordenadas y comprimidas.



## Otros formatos. NOOK

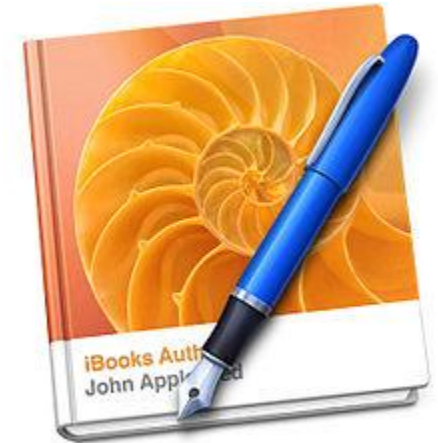
- NOOK/ePib, para libros-e para niños (fixed-layout).
- Propietario de Barnes and Noble para su plataforma de comercialización NOOK.
- Los contenidos se basan en .PDF (no HTML/CSS).
- Software lector disponible para tablet PC y otros dispositivos móviles.
- Creación restringida a editores que comercializan vía Barnes and Noble.

**BARNES & NOBLE**  
BN.com



# Otros formatos. Apple iBooks Author

- Basado en EPUB (HTML, CSS).
- Añade características que dependen del SW lector de Apple para obtener una funcionalidad completa “multi-touch”.
- Destinado a libros “complejos” en cuanto a presentación e interactividad.
- Extensiones para tipos de letras, ecuaciones matemáticas, y mayor interactividad.
- Comprometen compatibilidad con EPUB.
- Se generan vía Apple iBooks Author.
- Destinado a dispositivos apple.



[https://support.apple.com/kb/PH2789?locale=es\\_ES](https://support.apple.com/kb/PH2789?locale=es_ES)

# Formatos y preservación digital

- *“For ebook markup, most publishers currently use either the NCBI Book Tag Set or a proprietary format [...] while their scholarly books are on Amazon in a MOBI-based format, publishers tend to deliver the XML+PDF version of the ebooks to preservation institutions”.*
- Menciona el uso de PDF con metadatos ONIX:  
*“very few scholarly books are delivered to preservation institutions as full-text XML at this time”.*

# Formatos y preservación digital

- *“There is a growing consensus for the desirability of converging on standard open preservation formats for eBooks: in particular EPUB3 and NCBI BTS”*
- *“EPUB meets many of the requirements for a preservation-robust format”*
- *“Amazon Kindle formats dominate in the market ..., the proprietary nature of those formats and DRM comprise a long-term preservation risk”.*

.

# Conclusiones

- XML ha estado presente desde el inicio de las iniciativas libro-e, siendo OEB la especificación más representativa.
- XML se ha utilizado para:
  - Gestionar metadatos descriptivos y estructurales.
  - “Indirectamente”, para la codificación de contenidos (XHTML).
- La actual EPUB comparte las características de OEB respecto al uso de XML.
- Otros formatos (iBooks, KF8, FB2) hacen uso de XML (los dos primeros, de forma “similar a EPUB)

.

# Conclusiones

- Como “aplicación” de XML, EPUB comparte las ventajas de este:
  - Independencia de fabricantes (preservación, mercado más amplio)
  - Posibilidad de “reformatear” y “reutilizar” el contenido para distintas aplicaciones, contenidos y dispositivos.
  - Mayor número de herramientas para su procesamiento.
  - Su carácter abierto le da unas mayores garantías de “pervivir” en el futuro.
- Frente a PDF, son documentos más fáciles de procesar.

.

# Conclusiones

- ¿Pero se cumplen todas las posibilidades de XML?
  - EPUB se usa fundamentalmente como un formato de distribución.
  - Las mayores posibilidades de XML están en la explotación de los repositorios de contenidos, a partir de los cuales se pueden generar distintos “productos”, entre ellos EPUB.
- En el contexto de la autoedición/repositorios institucionales, ¿se está considerando este formato?
- ¿Deben las instituciones fomentar el uso de estos formatos alternativos a PDF?

.

# Conclusiones

- La adopción de formatos estructurados implica costes, y por lo tanto riesgos.
- Existen muchas oportunidades de automatización basadas en XML y en las especificaciones relacionadas.
- La definición de una estrategia implica:
  - Análisis de contenidos actuales y necesidades de mercado (trade-off).
  - Selección de formatos y estructuras de almacenamiento flexibles, que permitan reformatear fácilmente los contenidos.
  - Conversión y etiquetado de contenidos.
  - Definición de procesos de conversión o transformación de contenidos a múltiples formatos de salida.
  - Definición - opcional - de procesos de “enriquecimiento” de contenidos.