

# **Book Expo America / IDPF: Managing the Digital Warehouse**

Issues in the Production and Delivery  
of Digital Reflowable Books

*Jonathan Hevenstone*      *President and General Manager*  
*Publishing Dimensions*



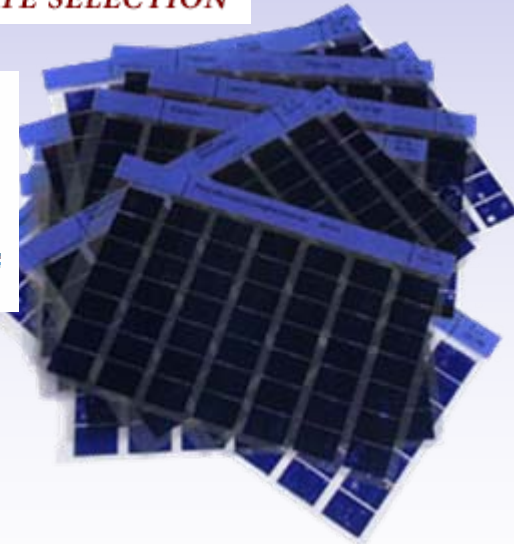
*...to...*



Book of the Month Club  
**BOMC**  
ALTERNATE SELECTION



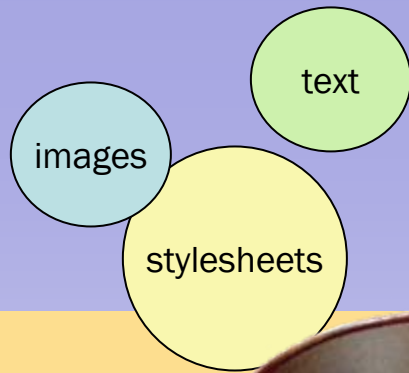
book.pdf



Production files repurposed to create ancillary products

PUBLISHING  
Dimensions

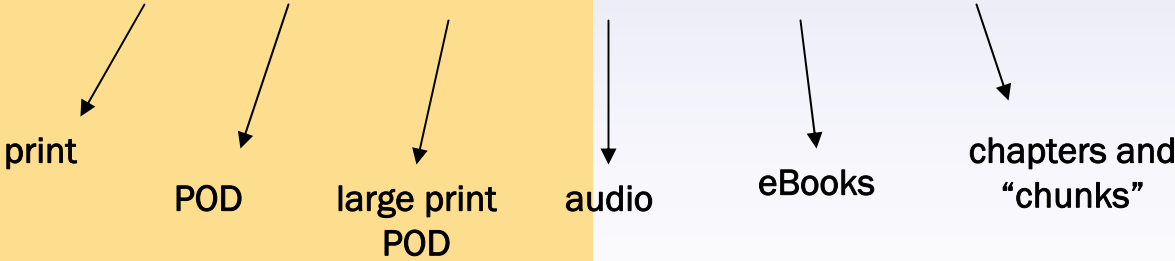
*...to*



metadata

rights information

<XML composition



The "book" as an abstract resource



# The future:

## The “book” as an abstract resource

- Open-ended, ongoing process of product and tool development
- Write new scripts/routines to generate product for new platforms and distribution models
- Fully automated generation and distribution of known products
- Data structured and ready for development of products and services beyond the horizon

# How we'll get there

- More XML up front—in the composition process, and even in the content development process
- Increased (and increasingly effective) adoption and leveraging of Digital Asset Management systems
- Increased standards development and integration for content tagging, metadata, DRM, ...

# Why we're not there yet:

## Issues in eBook production and delivery

1. **Special formatting and workarounds** are required to gear tagged content for each different Reading System
2. Complicated, **nonstandard** distribution procedures
3. **Metadata discontinuities** across publishers and retailers

# 1. Special formatting required for each Reading System

Differences across eBook formats mean OEB files must be processed differently for each Reading System.

Issues include:

- Entity support
- Table handling
- Font size settings
- Image format and filesize support
- Other tags and formatting elements

# Example: numbering in ordered list

Format 1: `<ul class="unOrderedList0">`  
`<li class="listItem">1. possibility of secret ploy</li>`  
`<li class="listItem">2. safeguard against exposure</li>`  
`</ul>`

Format 2: `<blockquote>`  
`<ol class="unOrderedList0">`  
`<li class="listItem">1. possibility of secret ploy</li>`  
`<li class="listItem">2. safeguard against exposure </li>`  
`</ol>`  
`</blockquote>`

Format 3: `<div class="bulletList">`  
`<p class="listItem">1. possibility of secret ploy</p>`  
`<p class="listItem">2. safeguard against exposure </p>`  
`</div>`

Format 4: `\t1. possibility of secret ploy\t`  
`\t2. safeguard against exposure\t`

## 2. Distribution procedures not standardized

MATERIALS READY FOR DELIVERY			
	<u>metadata</u>	<u>mktg images</u>	<u>eBook formats</u>
<b>RETAILER 1</b> ▶	Email metadata spreadsheet, then send content files	Rename mktg images & eBook formats per specs and deliver by FTP	
<b>RETAILER 2</b> ▶	Upload metadata, mktg images, and eBook files via UI (single operation)		
<b>RETAILER 3</b> ▶	Upload metadata, mktg images, and eBook files via UI (single operation)		
<b>RETAILER 4</b> ▶	Email metadata spreadsheet, then send content files	Rename mktg images & eBook formats per specs and deliver by FTP	
<b>RETAILER 5</b> ▶	FTP metadata XML & spreadsheet, then notify by email	Deliver packaged eBook files & mktg images via FTP	

### 3. Metadata discontinuities across publishers and retailers

- Publisher may provide ONIX, Excel, email, or Word
- Publishers' metadata often tailored to in-house needs
- Different retailers require different metadata
- Retailers don't necessarily accommodate metadata that publishers deem important. Publishers don't necessarily maintain metadata that retailers deem important.
- Retailers may require use of web interface for uploading eBook and metadata files

# Comparison of metadata from four publishers

<b>Fields:</b>	<b>Pub 1 (Excel)</b>	<b>Pub 2 (ONIX)</b>	<b>Pub 3 (Excel)</b>	<b>Pub 4 (Excel)</b>
ISBN				
Title				
Subtitle				
Volume #				
Edition #				
Series Title				
Number within Series				
Contributor 1				
Contributor Function 1				
Contributor 2				
Function 2				
Language				
Binding				
E-Book Format Code				
Illustrated (Y or N)				
Pub Date (YYYYMMDD)	On Sale			
On Sale Date (YYYYMMDD)				
Status				
Publisher				
Imprint				

(more...)

# The IDPF's new eBook specifications

- OPS 2.0 and OPF 2.0
- OCF

A step toward automated eBook production and a single delivery method.

# What the eBook world needs now is...

1. ...for all Reading Systems to support a uniform entity list (Unicode)
2. ...for all Reading Systems to support all image formats included in the OEB specification
3. ...for all eBook formats and platform rendering applications to fully support CSS 2.0 (included in OEB 1.2 and OPS)

**The new OPS specification will enable validation, a key step toward reaching these goals.**

# Other formatting benefits of OPS

- Font embedding
- Support for SVG (Scalable Vector Graphics)
- Incorporates the DAISY DTBook (Digital Talking Book) specification alongside XHTML as a Preferred Vocabulary
- NCX for enhanced navigation and accessibility

# Benefits of OCF as a distribution format

- Standard way to deliver content to retail channels
- Single file based on common ZIP compression
- OEB and OCF spec can evolve to incorporate the delivery of metadata, marketing images, and other resources—  
incentivizing standardization

## Next steps...

- Implementation by reading systems
- Adoption by publishers
- Support from service providers
- Development of production tools
- Development of validation tools
- Gather feedback toward future versions of these specifications