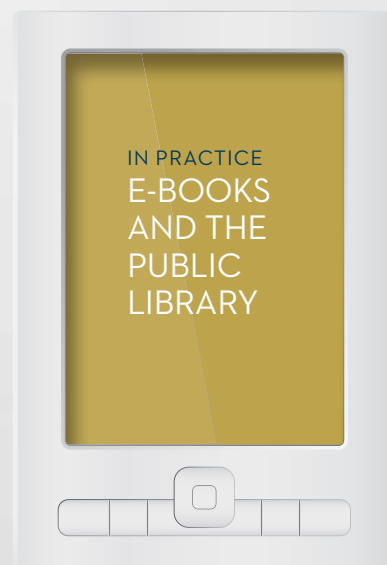
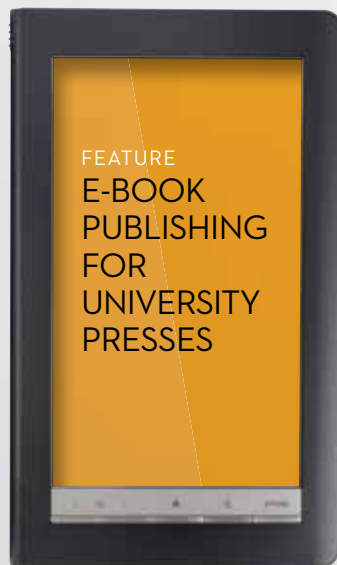


ISQ

INFORMATION STANDARDS QUARTERLY

SPRING 2011 | VOL 23 | ISSUE 2 | ISSN 1041-0031

SPECIAL EDITION: VIEWS OF THE E-BOOK RENAISSANCE



2011 NISO EDUCATIONAL EVENTS

www.niso.org/news/events 

JUNE

23–28 NISO at ALA Annual
(New Orleans, LA)

24 NISO/BISG Forum:
The Changing Standards
Landscape

26 NISO Update

JULY

No events are held this month.

AUGUST

10 Managing Physical Storage
(NISO Webinar)

SEPTEMBER

14 Preserving Digital Content
(NISO Webinar)

28 ROI in Linking the Semantic
Web (NISO Webinar)

OCTOBER

NISO Two-Part Webinar: Data
(NISO Webinar)

12 Data: Supplemental
Materials (Part 1)

19 Data: Technical
Management (Part 2)

24–25 The E-books
Environment
(NISO Two-Day Forum
Washington, DC)

NOVEMBER

9 New Discovery Tools
(NISO Webinar)

DECEMBER

14 Assessment Metrics
(NISO Webinar)

NISO OPEN TELECONFERENCES

Join us each month for NISO's Open Teleconferences—an ongoing series of calls held on the second Monday of each month as a way to keep the community **informed** of NISO's activities. The calls also provide an opportunity for you to give feedback to NISO on our activities or make suggestions about new activities we should be engaging in.

The call is free and anyone is welcome to participate in the conversation.
All calls are held from 3:00 – 4:00 p.m. Eastern time.

Webinar Subscription
Package Discounts
Buy all 14 for the price of 7!
Buy 4 and get 3 free.

ISQ

SPRING 2011 | VOL 23 | ISSUE 2 | ISSN 1041-0031

NISO

How the information world
CONNECTS

INFORMATION STANDARDS QUARTERLY (ISQ) is a publication by the National Information Standards Organization (NISO). ISQ is NISO's print and electronic magazine for communicating standards-based technology and best practices in library, publishing, and information technology, particularly where these three areas overlap. ISQ reports on the progress of active developments and also on implementations, case studies, and best practices that show potentially replicable efforts.

NISO MANAGING DIRECTOR and ISQ PUBLISHER | Todd Carpenter

ISQ MANAGING EDITOR | Cynthia Hodgson

NISO ASSOCIATE DIRECTOR FOR PROGRAMS | Nettie Lagace

NISO BUSINESS DEVELOPMENT & OPERATIONS MANAGER | Victoria Kinnear

DESIGN | B. Creative Group, Inc.

ISQ Board Members

Marshall Breeding, *Vanderbilt University*

Priscilla Caplan, *Florida Center for Library Automation*

Helen Henderson, *Ringgold, Inc.*

Norm Medeiros, *Haverford College*

Peter Murray, *OhioLINK*

Andrew Pace, *OCLC*

Kristen Ratan, *HighWire Press*

Kevin Thomas, *Wolters Kluwer Health*

Kate Wittenberg, *Ithaka*

2011 AD RATES

	1 ISSUE	2-3 ISSUES	4 ISSUES
Full page (8.5" x 11")	\$375	\$350	\$325
Half page (4.25" x 5.5")	\$250	\$225	\$200
Back cover (8.5" x 11")	\$700	\$600	\$550
Inside Cover Front (8.5" x 11")	\$500	\$450	\$400
Inside Cover Back (8.5" x 11")	\$500	\$450	\$400

For more information on advertising, visit www.niso.org/publications/isq

CONTENTS

3 From Guest Content Editor

FE 4

FEATURE

- 4 EPUB 3: Not Your Father's EPUB
- 12 Ten Questions and Tentative Answers about the State of E-book Publishing for University Presses

IP 16

IN PRACTICE

- 16 E-books and the Public Library: the Evansville Vanderburgh Public Library Experience

OP 21

OPINION

- 21 Drinking the E-book Kool-Aid in a Large Academic Library

SP 25

SPOTLIGHT

- 25 The Challenge for Standards in the E-book Supply Chain
- 31 E-books on EBSCOhost®: Combining NetLibrary™ E-books with the EBSCOhost Platform

NR 35

NISO REPORTS

- 35 The Evolution of Accessible Publishing: Revising the Z39.86 DAISY Standard
- 40 NISO E-books SIG

NW 42

NOTEWORTHY

SS 48

STATE OF THE STANDARDS



ATTEND ONE FREE!
REGISTER FOR ALL THREE DCMI
WEBINARS AND ATTEND NISO'S
SEPTEMBER 28 WEBINAR ON
SEMANTIC WEB ROI FOR FREE,
EQUIVALENT TO A
25% DISCOUNT ON THE
FOUR EVENTS.

JOINT 2011 WEBINARS

The National Information Standards Organization (NISO) and the Dublin Core Metadata Initiative (DCMI) are continuing their educational partnership in 2011 with three joint webinars on topics related to metadata, linked data, and RDA. This partnership builds on the highly successful September 2010 NISO/DCMI webinar that drew over 350 people from more than 100 sites.

MARCH 16

Metadata Harmonization:
Making Standards Work
Together

AUGUST 24

International Bibliographic
Standards, Linked Data,
and the Impact on Library
Cataloging

NOVEMBER 16

The RDA Vocabularies:
Implementation,
Extension, and Mapping

Each webinar will take place from **1:00-2:30 p.m.** (Eastern time) • *Registrants will receive access for one year to the recorded webinar.*



October Ivins

FROM THE GUEST CONTENT EDITOR

Views of the E-book Renaissance

Those of us who participated in the first round of e-book development (1998 to 2001) understand too well how the lack of standards hinders not only development and investment, but sales and usage. Our goal for this issue of ISQ is to present an overview of the status of e-books from multiple perspectives—publishers and other content producers, librarians, and the many vendors who support their creation, management, sales, and distribution. Not coincidentally, it also illustrates the scope of the NISO community.

The Open E-book Forum, an industry group focused on trade publishing, created the OEPS (Open e-book Publication Structure) in 1999 and continued to enhance and maintain it. In 2005, the OEB became the IDPF (International Digital Publishing Forum) to reflect interest in all types of digital publications and released the initial EPUB specification in 2007. Last year they welcomed collaboration with scholarly and academic publishers and vendors on EPUB 3.

The first feature article, by BILL KASDORF, is a lively explanation of the process of adapting EPUB 2.0.1, a “flowable” format created to display fiction on multiple devices, to EPUB 3, a format that supports much of the scholarly apparatus required for academic publications. The new specification was approved by the Board in May after a comments period and is expected to become a final IDPF Recommended Specification later this summer. University press director MARLIE WASSERMAN contributes a second feature, answering 10 ill-informed questions or assumptions about e-books that academic publishers must address to succeed as e-book producers.

The perspectives of both public and academic librarians are displayed in two submissions. MARCIA LEARNED AU and MOLLIE PHARO present a cogent case study of a medium-sized public library that has built an e-book collection over 11 years. **Au and Pharo observe that their downloads from Overdrive increased 500% from Jan. 2010 to Jan. 2011, and 800% in the last quarter of 2010 alone.** Since PDFs are still the preferred format for academic books, and laptops and desktops the preferred devices for reading them,


their experience provides valuable insights about device use and training, policy development, and more for public librarians. WENDY SHELBURNE contributes an entertaining opinion piece outlining her love/hate relationship with e-books as an academic librarian. The role for more and better standards is apparent in both of these contributions.

The standards spotlight is a gem. MARK BIDE offers a clear overview of the various standards initiatives EDItEUR is developing, and how they interact with other standards. His descriptions of the challenges of standardizing e-book metadata makes the task seem even more difficult than one imagined. For the NISO member report, we invited EBSCO to report on their progress in revamping NetLibrary, a legacy e-book system in need of both technological and business model revision. MICHAEL GORRELL explains EBSCO’s plans, and provides a timetable for their product redevelopment.

The first NISO report is an excellent description by MATT GARRISH and MARKUS GYLLING on the evolution of the DAISY standard to make e-books accessible for visually impaired users. The last article is a discussion by TODD CARPENTER of the newly created NISO E-books Special Interest Group. E-book interest is not new for NISO, who with OCLC commissioned a valuable study in 2009 by JUDY LUTHER, *Streamlining E-book Metadata Workflow*. We can look forward to the continued cooperation among standards groups that e-books will need to reach their full potential.

October Ivins

October Ivins | Guest Content Editor

 Streamlining E-book Metadata Workflow:
[www.niso.org/publications/white_papers/
 StreamlineBookMetadataWorkflowWhitePaper.pdf](http://www.niso.org/publications/white_papers/StreamlineBookMetadataWorkflowWhitePaper.pdf)

EPUB 3

(Not Your Father's EPUB)

OPENING PANDORA'S BOX IN THE WORLD OF E-BOOKS

BILL KASDORF



HTML 5

With all its new capabilities—handling rich media, complex layouts, scripting, global language support, MathML, synchronizing text and audio, and a host of other new features—EPUB 3, the new generation of the EPUB specification just issued by the IDPF (the International Digital Publishing Forum), may seem to be opening Pandora’s Box in the world of e-books.

I’d rather make the case that it’s trying to keep the lid on it—or at least trying to open the lid carefully, in the hope that all the creatures bursting out can be made to behave in a civilized way. That may seem to be a vain hope, but it’s a noble one. And I’m betting it will be successful.

Fundamentally, the revision of the EPUB specification was a response to the seemingly out-of-control pace of change in the world of e-publishing. Publishers are chafing at the limitations of today’s two fundamental e-book formats, PDF and EPUB 2.0.1, as advances in the wider world of the Web, the proliferation of new devices, and, most of all, the ubiquity of smartphones create both demand for more sophisticated functionality and impatience with solutions that fall short.

Amazon didn’t name its e-reader “Kindle” by accident: Jeff Bezos wanted to light a fire, and he did. The stunning success of the iPhone was succeeded by the even more stunning success of the iPad. And in the meantime, we’ve all become Google-eyed. Suddenly, everybody wants to get everything (information, entertainment, instruction) in

CONTINUED »



It's often a revelation to find that an EPUB file is really a zip file: if you change the .epub extension to .zip, you can see all the goodies inside—metadata, XML files, images, and so forth. (Go ahead, try it.) But it's not *just* a zip file, it's a particular *kind* of zip file following very particular specifications—the EPUB specifications.

whatever form they prefer (text, video, audio) on whatever device they happen to be holding (laptop, e-reader, tablet, smartphone), whenever they want it, and wherever they happen to be. Overnight, our information landscape seems to have turned into another Pandora: James Cameron's bewildering *Avatar* landscape.

Most shocking of all, this is not a futuristic fever-dream of a bunch of techies who've had too much Mountain Dew. This is not a movie. This is real. We really *can* start reading a book on a laptop in the office and pick up where we left off on a phone on our way home. We really *can* click on the name of a baseball player in a magazine on a tablet and see his stats pop up or watch him hit that pennant-winning homer. We really *can* see and hear Dr. Martin Luther King's "I Have a Dream" speech in an encyclopedia. We really *can* do a self-test in a textbook and get an instant grade (and receive guided learning based on how well we did).

We even take these things for granted now. That's because the Web already enables all those things. The good news: the Web lets us do an amazing number of really cool things. The bad news: the Web lets us do them in lots of different ways, using lots of different technologies. Many of which are proprietary. Many of which are incompatible with each other.

Why isn't this more of a problem on the Web? There are two main reasons. First, we have a choice of browsers, and they're free and frequently updated, so if the browser we're using isn't able to handle something, we can often find another one that can. Second, "online" is a two-way street: the sender of the content can detect important information about the capabilities of the recipient, and can adjust what it's sending and how it's sending it. EPUBs can't always do that. Fundamentally, EPUB is a "package" masquerading as a file format, designed to be accessible and fully functional *offline*.

It's often a revelation to find that an EPUB file is really a zip file: if you change the .epub extension to .zip, you can see all the goodies inside—metadata, XML files, images, and so forth. (Go ahead, try it.) But it's not *just* a zip file, it's a particular *kind* of zip file following very particular specifications—the EPUB specifications. Specifications for how the text content is marked up in XML. Specifications for what metadata must be there, and how it must be expressed. Specifications for what other file types can be included. Specifications for the file types that conformant reading systems must be able to handle. Specifications for how all the pieces are organized—and how to *say* how the pieces are organized. (And lots more.) The Web doesn't have to do any of that.

Although these specifications may seem to be making things unduly complicated, they actually make things much easier. Easier for the publisher of the EPUB, who—faced with our Pandora's swarm of choices and options—has a clear path to consistency. Easier, too, for the maker of the reading systems that need to receive and render the EPUBs: instead of needing to be able to accommodate that multitude of options—or, worse, *not* to accommodate some of them—the system can know what it will get in an EPUB, how it can find the pieces it needs, and what it needs to do with them. All of which, ideally, makes it easier for the consumer as well.

First, a word about our sponsor

You may have noted that I've been careful to use the word "specification," not "standard," when referring to EPUB. The reason is important. It is, of course, intended to provide a standard way in which to interchange and deliver reflowable content to reading systems. ("Reading systems" is also a carefully chosen term.



EPUB 3 is designed to accommodate textbooks, scholarly and STM monographs, and technical manuals as well as non-book content like magazines, newspapers, journals, white papers, and corporate documents: anything anybody would find it useful to package as an EPUB and interchange or deliver through any of the rapidly proliferating choices of reading systems.

EPUBs aren't just for handheld e-reading devices; they're for laptops and desktop computers and text-to-speech reading systems for the print disabled—any e-reading environment, including ones not invented yet—as well.) For all practical purposes, it's a standard, just as the current EPUB 2.0.1 specification is. While it is not yet an International Standard from a de jure standards body like NISO, it is a standard similar to those from the W3C, the World Wide Web Consortium, the keeper of the XML family of standards and most standards fundamental to the Web.

The IDPF, the International Digital Publishing Forum—the body that maintains EPUB—is a not-for-profit trade association of over 200 members from over twenty countries that represents a broad cross-section of the publishing ecosystem: publishers, technology companies, device manufacturers, and others, commercial and non-commercial, with a shared interest in fostering a robust and well functioning e-publishing environment. Its EPUB 3 working group is an extraordinarily large one—170-some members, with additional invited experts—representing that full spectrum.

It was a veritable “Peaceable Kingdom” of competitors collaborating: publishers of all types, from some of the largest commercial and nonprofit trade, scientific, scholarly, and educational publishers to the smallest; makers of e-book devices (both Apple and Sony were very active members); e-book distributors and retailers; technologists from companies like Adobe and Google to small consultancies and individual programmers; people from standards organizations like DAISY, NISO, EDItEUR, BISG, and IDEAlliance (though none *formally* representing those organizations, I should make clear); and many others, such as librarians, service providers, and other interested parties. It was also genuinely international, with particularly active and valuable participation from Japan and elsewhere in East Asia.

These folks obviously didn't always agree on everything, but they did agree on the group's fundamental mission, and they worked conscientiously (and *hard*) to come up with the best possible result. All of this work was done in a totally transparent fashion, on an open listserv and a publicly accessible wiki, with a strict avoidance of proprietary issues and intellectual property claims. There were clear mandates for what was to be accomplished and a formal process for accomplishing it. The result is an open standard, based as much as possible on open standards, that addresses real-world needs in realistic ways.

It's called “EPUB” not “EBOOK”

One of the most important goals for EPUB 3 was to accommodate a much broader range of content than EPUB 2.0.1 did. The IDPF originated mainly in trade publishing, and the earlier generations of EPUB reflected that. In contrast, EPUB 3 is designed to accommodate textbooks, scholarly and STM monographs, and technical manuals as well as non-book content like magazines, newspapers, journals, white papers, and corporate documents: anything anybody would find it useful to package as an EPUB and interchange or deliver through any of the rapidly proliferating choices of reading systems.

EPUB 3 was also designed to accommodate a much broader range of *types* of content. No longer primarily for simple text-and-image content, it now provides a practical solution for incorporating multimedia content like audio and video, as well as animations and other scripted functionality. In keeping with its international mission, it provides global language support. It provides much more sophisticated typographic and layout capabilities (this is especially important to magazine and textbook publishers). It accommodates much more extensive metadata, at all levels, from the package to the paragraph. And in keeping with one

CONTINUED »

of its most important mandates, it is designed to enable and facilitate conformance with standards for accessibility. (These aspects of EPUB 3 will all be discussed in more detail below.)

Creating a specification out of squishy standards

A dilemma facing the EPUB 3 Working Group from the outset was that some of the standards that were obvious candidates for inclusion in the EPUB 3 spec are not yet fully and formally final. (The technology folks call them “squishy” and “not fully baked,” in case you want to know the true technical terminology.) The most important of these were HTML5 and CSS3.

I mentioned that EPUB 3 was created in response to our rapidly changing e-publishing environment. I should state that more strongly: EPUB 3 was created in a mad dash to get ahead of the rapidly changing technological developments before it was too late to make a difference. To put it bluntly: *people are already doing all this stuff*. And although one of the problems EPUB 3 was created to address is the chaotic assortment of technologies and techniques being used today, the reality is that there are already some clear best practices. It would be foolhardy for EPUB 3 to try to force people in directions in which they clearly don't want to go, or which are recognized as bad ways to go, or for which it is already too late.

So, for example, while it prompted a bit of discussion (these working groups are, after all, the descendants of college debating societies and dorm room arguments), the decision on font formats was really pretty clear cut: OpenType and WOFF. OpenType is clearly the right font format from the print publishing world, from which lots of EPUB content comes; WOFF is clearly the dominant font format in the Web publishing world; and you can't easily make one conform to the other. Ergo, EPUBs have to use either OpenType or WOFF fonts, and reading systems have to support both. End of discussion.

The core standards at issue in this regard were HTML5 and CSS3. These are extremely extensive and fundamental standards. HTML is the language of the Web and CSS (Cascading Style Sheets) is the way the Web is mainly rendered (or should be). Their latest incarnations, HTML5 and CSS3, can be thought of as the next generation of the Web, taking shape in front of our very eyes. People already use them to do cool things with typography and layout, interactivity, animations, rich media. Browsers already implement them. Yet they are still both works in process. While they are both in a sense “modular,” important modules are nowhere close to being finalized. HTML5 is not expected to be a formal Recommendation of the W3C until 2014. There will be millions of EPUBs created between now

and 2014. EPUB 3 can't wait. So the dilemma boils down to this: EPUB 3 really *must* be based on HTML5 and CSS3, and yet they're not really finished standards.

The resolution, I think, was eminently reasonable. The EPUB 3 Working Group elected to selectively specify those modules of HTML5 and CSS3 that either (1) are in fact considered finished, for all practical purposes, or (2) are essential to an EPUB 3 requirement and are close enough to resolution that they are reasonably safe to use. Moreover, the EPUB 3 spec attaches a “warning label” to the latter: if at some point the HTML5 or CSS3 spec changes from what EPUB 3 is specifying, EPUB 3 makes the commitment to change along with it, so that EPUB 3 will stay in synch with HTML5 and CSS3. This approach is realistic, practical, and not as risky as it appears to be. For example, EPUB 3's use of CSS is really still almost entirely based on the existing CSS 2.1 specification; it just brings in certain modules from CSS 3 that are needed to accomplish certain things that CSS 2.1 does not address.

Markup and metadata

Although the interest in EPUB 3 is understandably focused on all the new capabilities it offers, it is important to understand that it is fully backwards compatible with EPUB 2.0.1. That means that all EPUB 3 conformant reading systems *must* render EPUB 2.0.1 publications properly. It also means that most of the new aspects of EPUB 3 are optional.

One very important change concerns the text markup vocabulary. EPUB 2.0.1 provided two vocabularies: XHTML (which was used by the vast majority of EPUBs) and DTBook, a vocabulary published by the DAISY Consortium for accessibility purposes. DTBook has been eliminated from the EPUB spec because the DAISY Consortium decided to work with the IDPF to enable EPUB 3 to become the distribution format for accessible content, rather than requiring a separate model as they had before. Thus DTBook is being phased out as a delivery format by both DAISY and EPUB, leaving the basic markup vocabulary for textual content in EPUB as XHTML—the vocabulary used by the overwhelming majority of EPUBs created so far. XHTML5 (the XML expression of HTML5) as used in EPUB 3 provides *additional* vocabulary features but does not change the basic XHTML vocabulary used by the previous spec.

There has been one change that's particularly important to STM publishers: MathML, the standard for representing mathematics in XML, is now a “first class citizen” in EPUB 3. MathML provides both semantic and presentational markup, the former concerned with what math expressions mean and the latter with what they look like. It is only the *presentational* aspects of MathML that EPUB 3 reading systems are required to support.

The metadata capabilities in EPUB 3 are also dramatically expanded from EPUB 2.0.1. While there are still only three required elements—`dc:identifier`, `dc:language`, and `dc:title`—EPUB 3 enables publishers to include many more identifiers, versions of titles (for example a title for sorting purposes or a short title), and many more elements of metadata, along with specifying the scheme that defines them (e.g., MARC 21, ONIX 3.0, PRISM).

There is a very basic set of standard metadata terms “built in” to EPUB 3 that can be used without a prefix; all other metadata requires a “profile” to be declared and a prefix to be used on those metadata elements. This enables content creators from many different interest groups with specific metadata practices to incorporate their own metadata vocabularies within an EPUB. (It’s a quasi-namespaces approach, but it does not require resolution to a metadata authority. However, it does require *identification* of the appropriate metadata authority.)

In addition, EPUB 3 enables content creators to associate metadata with EPUBs at all levels. Previously, it was only possible to associate metadata at the package level. Now, much expanded metadata can be associated with the package as a whole, with a component of the package, or even embedded right in the content markup itself, down to the paragraph level. (Phrase-level metadata can also be done.) Two of the special metadata attributes specified by EPUB 3 are `epub:type`, for adding semantic information to markup, and `epub:trigger`, to launch a multimedia or scripted function.

Finally, it is also possible to provide what is referred to as “external” metadata. This can either be a file of metadata included within the EPUB package (for example, a MARC record or an ONIX file) or pointed to via a link (which of course will only work in an online environment).

All of the metadata in EPUB 3 is expressed in very standard, widely used ways. Most is based on Dublin Core (with a preference for DCTERMS); other metadata is added using very simple features taken from RDFa 1.1. The goal was to provide a metadata mechanism that would be extremely easy to implement, even for nontechnical content providers, while accommodating the rich metadata that is becoming an ever more essential part of the information ecosystem.

Taking advantage of our new real estate

It’s a little ironic how much influence the emergence of tablets has had on the e-publishing landscape. Although there are many things to love about tablets—they are unquestionably already an indispensable component of our ecosystem—what seems at first to be their salient feature, the amount of real estate they offer in which to render content, has been there all along in desktop and laptop computers. (Of course it’s how that feature plays with all their other features—including their portability, their gesture-based interface, and their ability to be both “Web” and “not-Web”—that makes all the difference.)

It’s important to realize that what is most important about tablets from the EPUB point of view is that they are one mode among many in which to render EPUB content. They exist as a component of an ecosystem that does still include laptop and desktop computers, along with handheld reading devices and smartphones. Unlike PDF, which is a fixed-page format that locks in everything about the page (in fact, it is the stability of that format across media, from print to online, that is PDF’s greatest virtue), EPUB is all about *reflowable content*.

EPUB 3 provides the capability to design rich layouts like those common to magazines and textbooks—such as multiple columns (with hyphenation) whose



“Fixed Page” Layout

Through a function called “media queries” (the EPUB basically asks “where am I?”), different style sheets can be used to produce, for example, a two-page spread on a tablet held in landscape mode, a one-page two-column layout when that tablet is turned to portrait mode, and a single column format on a mobile phone, all from the same XHTML5 file. This enables a type of “fixed page” layout—control of the content page-by-page—while still enabling reflow.

CONTINUED »

CORE MEDIA TYPES



EPUB CONTENT DOCUMENTS

XHTML
SVG

FONTS

OpenType
WOFF

GRAPHICS

.jpg .gif
.png .svg

AUDIO

MP3
MP4 AAC LC

EPUB 3 specifies certain file types as Core Media Types. Reading systems must be able to render these properly, if they offer the functionality provided by a given Core Media Type. If a content provider uses a file that is not a Core Media Type, it must provide a fallback to a file that is a Core Media Type. EPUB 3 Core Media Types include those shown above.

CONTINUED »

text flows around images and sidebars—while enabling the design to adapt to the real estate available to it. Through a function called “media queries” (the EPUB basically asks “where am I?”), different style sheets can be used to produce, for example, a two-page spread on a tablet held in landscape mode, a one-page two-column layout when that tablet is turned to portrait mode, and a single column format on a mobile phone, all from the same XHTML5 file. This enables a type of “fixed page” layout—control of the content page-by-page—while still enabling reflow. This will be the most important feature of EPUB 3 for many publishers.

This is only one example of the dramatically improved capability for control of graphic design offered by EPUB 3. It not only permits embedded fonts, it encourages them (and provides for what is called “font obfuscation” to prevent font piracy). This is important not only for publishers who want to maintain branding or a certain “look and feel” for their publications, but also for publishers of specialized content like technical or linguistic content, which requires special “glyphs” that are unavailable on standard fonts.

Layout issues are of particular concern to Asian publishers or others who use non-Latin alphabets. Few people realize that EPUB 2.0.1 permitted right-to-left text reading (required for languages like Hebrew and Arabic) because reading systems didn’t implement this capability. EPUB 3 goes much farther, allowing vertical writing as well. In addition, SVG (Scalable Vector Graphics) is a Core Media Type (see figure above) in EPUB 3. Think of SVG as PDF expressed as XML: images are captured as vector graphics that adapt to the size and resolution of the rendering environment. This means that even publications like manga and graphic novels can be delivered as EPUB. And the

EPUB 3 spec enables the publisher to specify reading order, as well, so that a book can be read from right to left and the first page of a spread can be understood to be on the right.

Rich media and scripting

The topic of rich media and scripting provides me an opportunity to address the “app” issue. The success of the iPad and the proliferation of competing tablets, along with the parallel proliferation of the iPhone and other smartphones—especially those created for the Android operating system—have made it possible for publishers to provide content that includes audio and video content as well as scripted behaviors ranging from simple animations to elaborate interactive functionality. Before EPUB 3, these were done through apps, the small single purpose applications that these technologies have made so popular. However, apps are an impractical way for most publishers to publish most of their content. They are specific to an operating system (an app for iOS won’t run on an Android device), they require programming, and they usually prove to be too expensive and time consuming to create for all but the most popular or high-priced products.

EPUB 3 enables all this to be done in a standard way that is device agnostic. Any reading system that is EPUB 3 conformant and which offers the ability to play video, audio, and scripts will properly render a conforming EPUB 3 publication. This is a huge benefit to content creators. No longer do they need to create different versions, with different specs and even different file formats, for different environments. Yet, in cases where it does still make sense to create an app—and there will be many such cases—doing so

is that much easier when it is done on the basis of an EPUB 3 in the first place. EPUB 3 helps answer the questions like: “Which format should my audio files be in?” and “Which scripting language should I use?” In an EPUB, an audio file will always be either MP3 or MP4 AAC LC (the latter because it’s required by the required MPEG 4 video format); if there is scripting, it will always be done using JavaScript. This makes things so much easier for content providers and reading systems alike!

Accessibility

EPUB 3 was created from the outset to address issues of accessibility, and experts from the DAISY Consortium have been instrumental in the development of the spec. An important benefit of this close collaboration is that in its update of DAISY (which is a NISO standard), the Consortium expects to specify EPUB 3 as the delivery format for DAISY. (See related article on page 35.) This means that in addition to using EPUB 3 to make files that will work on a host of reading systems, publishers can also use the same EPUB 3 files to deliver their content accessibly.

In EPUB 3, the navigational structure is specified not by the former proprietary format, but in an XHTML microformat. The XHTML5 markup of EPUB 3 content (with the addition of a new attribute, `epub:type`) can accommodate all the semantics necessary for accessibility. A new feature of EPUB, media overlays, enables the synchronization of text to audio, enabling print-disabled users to use the XHTML file to search and navigate its audio counterpart. And there are important text-to-speech features in the EPUB 3 spec, including PLS (the Pronunciation Lexicon Specification) and fine-grained pronunciation control via SSML (Speech Synthesis Markup Language).

Rapid EPUB 3 adoption is expected

Because EPUB 3 provides much needed clarity to our currently chaotic e-publishing environment, its adoption is expected to be swift. As soon as it is formally introduced, it is expected to be endorsed by major information industry organizations and adopted by major technology companies. There will be EPUB 3 reading systems available commercially before the end of 2011, and EPUB 3 is expected to be in wide use by 2012.

In the meantime, as soon as the specification is formally published, IDPF plans to publish extensive documentation, examples, and best practices to make it easy for publishers to incorporate EPUB 3 into their workflows, along with a validation mechanism to help ensure that EPUBs conform

properly to the specification.

While it’s clear that more work will need to be done to continue to advance the EPUB specification (which the IDPF plans to do in a modular fashion, rather than issuing future “monolithic” releases), the EPUB 3.0 spec is a major watershed. It will be the foundation on which our e-publishing ecosystem will be based for many years to come.

| FE | doi: 10.3789/isqv23n2.2011.02



BILL KASDORF -bkasdorf@apexcovantage.com is Vice President of Apex Content Solutions and General Editor of *The Columbia Guide to Digital Publishing*. He is a member of the IDPF’s EPUB 3 Working Group (and leads its Metadata Subgroup); BISG’s Content Structure Working Group (and chairs its Standards Survey Subgroup); and the IDEAlliance nextPub Working Group (chairing its EPUB-to-nextPub Mapping Committee). He is Past President of the Society for Scholarly Publishing and is a frequent speaker for publishing industry organizations.

CSS3

www.w3.org/TR/css3-text

Dublin Core Metadata Element Set

dublincore.org/documents/dces/

Dublin Core Metadata Terms [DCTERMS]

dublincore.org/documents/dcterm/

EPUB 3

dpf.org/epub/30

HTML5

www.w3.org/TR/html5/

International Digital Publishing Forum

www.idpf.org/

MathML

www.w3.org/TR/MathML3/

MPEG-4

mpeg.chiariglione.org/standards/mpeg-4/mpeg-4.htm

OpenType

www.microsoft.com/typography/otspec/

www.adobe.com/type/opentype/

Pronunciation Lexicon Specification (PLS)

www.w3.org/TR/pronunciation-lexicon/

RDFa 1.1 Primer

www.w3.org/TR/rdfa-primer/

Speech Synthesis Markup Language (SSML)

www.w3.org/TR/speech-synthesis/

Web Open Font Format (WOFF)

www.w3.org/TR/WOFF



RELEVANT
LINKS

10 Questions & TENTATIVE ANSWERS

THE STATE OF E-BOOK PUBLISHING FOR UNIVERSITY PRESSES

Fifteen years since we first heard the term e-book in publishing circles we have more questions than answers. As I look at e-book publishing from the perspective of a director at a medium-sized university press situated in a large, public, research-oriented university, I see many exciting opportunities but no clear roadmap to success. Most university presses have more of a will than a way, although their paths vary by the extent of their staffs and financial resources. Let's look at ten of the prevailing questions that are on the minds of either librarians, customers, or press directors and explore the current state of progress.

1

Can't publishers solve all their financial problems by simply switching from print books to e-books?

This is the mother of all myths, and it has had remarkable staying power. After all, won't publishers save money on paper, printing, binding, and warehousing? No matter how many times publishers explain the oversimplicity of that argument, it prevails. I have had long discussions with tax lawyers, accountants, and mathematicians, explaining the complexity of the finances, but to no avail. Here is one more stab. For many publishers—including university press publishers—paper, printing, binding, and warehousing make up a small percentage of the total cost of publication. Other costs remain constant, including what may be years spent enticing an author to publish with that press, travel to conferences to meet that author, time spent suggesting reorganization of the project, time and cost connected with peer reviews of a manuscript, copy editing, design, marketing, and the accounting work of royalty statements and vendor payments. Estimates vary, but in general the costs that are constant are likely to make up 85% of the total cost of publication, with the remaining 15% going to the costs that would go away in a purely electronic world. But other costs must be added for e-books. These new costs involve

personnel who modify digital files according to the multiple standards specified by multiple vendors, prepare metadata again according to multiple standards from multiple vendors, record micropayments that may be as low as 17 cents so they can be added into royalty statements, and set up archiving systems for digital files. Some medium-sized university presses have managed to hire content services system providers, also known as digital asset management systems, that carry out these functions, but at a high cost that is often the equivalent of one or two staff members. These expenses generally rise to about the same amount of money a publisher saves by not printing, binding, and warehousing. In short, I argue that there are many reasons to publish books electronically, but saving money is not one of them.

2

Shouldn't e-books cost the customer less than print books?

As noted above, the publisher is not really saving any money. In the meantime, what happens to sales revenue? Can the publisher charge the same price for the e-book as for the print book, in order to recover costs that might look a little different but still add up to the same amount? There are two schools of thought here. First, and alarmingly to the publishing world, we

see the position of many commercial e-vendors that the e-book should be priced lower than the print book, based on the false assumption that it has cost less money to produce. An extreme example of this was Amazon's original practice of charging \$9.99 for most Kindle editions. Second, and alarmingly for the library world, we see some vendors taking the position that e-books should be priced higher than print books because they can be shared more easily (in some arrangements) and because of added functionality (searching). Publishers must find a way to recover their non-reduced costs through the sale of e-books, but at prices libraries will consider justified.

3

Since university presses operate in a university community, can't they get help in migrating to an electronic model from other organizations in the university, particularly from libraries?

Many public universities are facing severe budget crises, which has been the case for years. At our university, each unit, including the press, must pay for its own tech support, hardware, and software. The server we use to archive our digital files was expensive, as was the tech support to configure it. More seriously, our library has lost staff and funding and despite good will, is in no position to help us with data conversion, metadata, or archiving, at least certainly not on the tight schedule required for most e-book initiatives. The library administrators have started to research and account for some of the costs associated with their own entrepreneurial efforts in the world of e-publishing and archiving, and fully realize these initiatives are not free. There are indeed a few university libraries that have taken a lead role in e-publishing, but these libraries represent the exceptions, not the rule.

4

Why don't university presses sell e-books themselves? What happened to the trend toward disaggregation?

Another unresolved issue concerns how many chefs stir the pot—also known as aggregation versus disaggregation. Many university presses of all sizes began by working with NetLibrary, which at first digitized our books for no cost. In that scenario, three parties participated: the publisher, the intermediary, and the library. Variations on that model continued with a host of additional vendors, including ebrary, Questia, MyiLibrary, EBSCO, and so on. This same pattern

holds for Amazon's Kindle books and Barnes & Noble's Nook books. But there are other experiments going on that have fewer parties. Tizra is a company that has licensed software to presses to enable those presses to sell their books on their own websites. The University of Chicago Press has been renting a sampling of its books on their website for a specified period of time, with no third-party intermediary. Direct selling enables a press to keep a higher percentage of sales revenue than is the case where there are intermediaries. These laudable experiments are not without problems. Direct selling requires presses to create additional marketing efforts that drive customers to their websites. Direct selling may not be convenient for students and scholars who want one-stop shopping if they are buying books from multiple presses. Another complication is that for a press to sell or rent on its own site requires either licensing software, which is an expense, or a substantial and high-powered staff of technically adept experts. There may also be complications with protecting charge numbers, unless the press has already been selling print books online over its own website as opposed to a system where it appears the press is doing that but a larger company is actually safeguarding the commercial data. In short, very few presses are able to sell directly to the customer. So although disaggregation is one trend, it is unclear whether it will gain momentum.

5

Won't e-books displace print books?

This myth is particularly puzzling to me because there is no single model of technological change in our daily lives. Perhaps iTunes has displaced records and digital pictures have replaced film, but television did not eradicate radio and DVDs did not eradicate movie going. For the foreseeable future, e-books will exist alongside print books. What this means for publishers is that we must have one foot in each world. For example, we must still spend about \$2000 designing and printing dust jackets, we must still produce print fliers as well as e-blasts, and we must still have physical books to exhibit at the annual meetings of professional societies. As new activities are added to our workload, no existing activities can disappear.

The Association of American University Presses (AAUP) has just circulated an insightful task force report entitled *Sustaining Scholarly Publishing: Business Models for University Presses*. It points out convincingly that university presses must have multiple business models operating at once, putting a burden on resources, and that it will be a long time before a single, standardized model emerges, if ever.

CONTINUED »

6

Why don't publishers prepare their digital files in XML?

Most university presses, with the exception of the largest and those with the most resources, have their digital files in PDF form because they cannot afford to migrate to XML. The cost of conversion has been coming down and may now be as low as \$100 per book. But again, this is a simplification of the problem. At Rutgers, we have about 700 titles in PDF. So we would be facing a \$70,000 expense, at a time when there is no extra money for initiatives. Even so, perhaps we could manage to do this if we could make a convincing case to ourselves and our university that we would recover that cost. There are certainly visionary people who can explain the added utility that comes with XML. But no one can assure us how that will convert into dollars, especially in the current economy. Moreover, if we listen to enough people, we hear that even XML may not be the ultimate solution. By waiting, we may find something even better. Finally, when we talk to most of the students and scholars who are likely to use our e-books, they feel that PDF, though not ideal, will give them most of the functionality they need.

7

Should we put our resources into selling individual titles, or should we bundle e-books into collections?

There is no easy answer to this question. From the point of view of advancing research, it is tempting to focus on selling collections of books, to facilitate thorough searches. Moreover, libraries seem comfortable with the notion of e-books as something akin to databases of knowledge. But librarians also like the ability to pick and choose rather than being told that if they want book A they must also buy book B, and they are experimenting with patron-driven acquisitions, which assumes sales of one title at a time. Individuals almost always prefer to buy single titles, that is if they have devices they can use as e-book readers. Yes, we can focus on both individual books and collections, but this involves dealing with endless vendors, who again have differing systems, requiring extra labor on the part of publishers.

8

Who are the players in the e-book market now?

There has been an enormous amount of activity in the past year or two as e-book vendors proliferate. In addition to the

vendors that university presses have worked with in the past—Questia, NetLibrary, ebrary, MyiLibrary, EBSCO, and so on—there are many newcomers, including Cambridge University Press and Oxford University Press, which are both distributing e-books for other publishers; JSTOR which plans to distribute e-books; and Project MUSE, which is about to launch an e-book venture called UPCC, or University Press Content Consortium. (This is the program Rutgers University Press has joined.) These initiatives are examples of the trend back to aggregation, as publishers realize the enormous resources required to sell effectively to libraries, domestic and international, and the advantages of collections for scholars doing research. Even many of the largest university presses are likely to participate in one or more of these collectivities. In addition, several of these initiatives have been developed with the benefit of surveys of librarians, and have attempted to be responsive to the needs of all parties in the scholarly communication ecosystem. Another factor that motivates at least some of these initiatives is the push toward partnerships that is rampant in many universities. The assumption is that we can all achieve economies of scale in a difficult economy by joining together.

9

As vendors proliferate, are libraries facing the prospect of buying the same content twice?

I and my fellow directors have, over the years, probably signed twenty to thirty contracts with e-vendors. We have been trained to hunt for the word exclusivity and if it is there, cross it out. As the field of e-books was developing, none of us knew which company would prove to be the best to work with, so we wanted to spread our valuable assets over multiple vendors, to see how business shook out. This may make sense when everyone is talking about individual titles, but for libraries, it may not make sense when the spotlight moves to collections. If one vendor offers a library a collection on U.S. history with 100 books, and another vendor offers libraries a similar collection, with 50% overlap, what is the result? Will the library buy one or both? Will the library need to put resources into assessing the level of overlap? Or do libraries welcome duplication so that they are not faced with what could be perceived as a monopoly? I have been part of the UPCC effort to bring university presses together to sell their e-books in subject collections to libraries. Although we recognized the advantages to libraries of insisting that if a publisher adds their books to one of our collections, they should not add them to another vendor's collections, we ultimately decided that publishers were not at a point where they were comfortable promising exclusivity.

10

Can't the problem of clearing rights for e-books be easily solved?

Publishers must ensure that they have permission to reprint whatever is in the content of their books. This commonly includes long quotes, lyrics, poetry, and images. For older books, rights may have been cleared for print editions, but not for electronic editions. I am not convinced that everyone in the world of scholarly communication comprehends the magnitude of this problem. For a press like ours, with about 3,000 titles in print, that requires looking at 3,000 paper files. Let's take a fictional book published in 1955 that incorporates illustrations and third-party text quoted from a journal or an earlier book. Did the press staff in 1955 keep meticulous records on what permissions were cleared, and whether they were cleared for all editions, in all formats, which is highly unlikely? If they were cleared for a print edition only, are the 1955 addresses of the rights holders still correct today so that we can ask for an extension of the grant of permissions to e-books? And if those addresses are valid, will those rights holders be willing to grant rights for an e-book without a fee? Then multiply these complexities by a factor of 3,000. The problem may not be this severe for those presses that give up on digitizing the longest of the long tail, or those presses that have kept accurate records for the last decade. But even under the best of circumstances, the workload is staggering. Few presses can accomplish rights clearance for e-books without a significant addition of staff. Some presses, especially those that are large and have substantial sales revenue, have hired consultants and freelancers to tackle the problem in a frenzy of activity. Other presses have done this gradually, with existing staff. Still others have not fully entered the world of e-books, largely because of this problem. At Rutgers, we have focused our attention on the more recent books, with the greatest sales potential.

The full burden of rights clearance need not fall on publishers alone. Many of us are putting that burden back on authors, asking them to clear rights for e-books at the same time they clear rights for print books. This practice appears to be sensible, but it results in a series of problems. First, many of the rights holders who respond to such requests come back to the author with a list of questions that don't seem to fit the current environment and are unanswerable. "How many copies will you be producing," they ask. Of course, we don't produce and we can't predict what is disseminated. "Will the e-book be combined with other material," they ask. Again, we cannot predict. After the questions come more irritations, when the fees charged for e-rights are often prohibitive, as though e-books are to be supplemental income rather than the substitute income that is more likely. University press authors who typically pay for permissions themselves are not pleased at either the questions or the fees.

In addition to workload and budget problems, rights issues limit the number of books a publisher can convert to e-books. Many university presses have published landmark anthologies that are widely used in classes. It would be prohibitive to even try to clear rights for such books, so we exclude them from consideration for e-books. Many of us publish imports from the U.K. and elsewhere, for which we do not have world rights. These must be excluded from certain ventures. Many of us publish art books, where again the fees would be prohibitive. For some publishers, these categories of books represent a small fraction of their lists, but for others, the percentage is significant.

The ten questions and tentative answers I have covered here point to the unsettled state of e-book publishing. We continue to deal with myths, unresolved issues, and unsettled standards. Using a life course perspective, I would say that e-books have grown out of infancy but are now in toddlerhood, struggling to talk and to stand. I look forward to the next stage. | FE | doi: 10.3789/isq23n2.2011.03



MARLIE WASSERMAN <marlie@rutgers.edu> is Director, Rutgers University Press (rutgerspress.rutgers.edu).

RELEVANT LINKS



Association of American University Presses (AAUP) » aaupnet.org

Cambridge University Press
www.cambridge.org/

ebrary » www.ebrary.com/

EBSCO Publishing
www.ebscohost.com/

JSTOR » www.jstor.org/

MyiLibrary » www.myilibrary.com/

NetLibrary » www.netlibrary.com/

Oxford University Press
www.oup.com/

Project MUSE » muse.jhu.edu/

Questia » www.questia.com

Rutgers University Press
rutgerspress.rutgers.edu/

Sustaining Scholarly Publishing: Business Models for University Presses

http://aaupnet.org/resources/reports/business_models/index.html
http://aaupnet.org/resources/reports/business_models/business_models2011.pdf

reports/business_models/
aaupbusinessmodels2011.pdf

Tizra » tizra.com/

Two University Press Ebook Initiatives Merge (University Press Ebook Consortium and Project MUSE Editions become the University Press Content Consortium)

www.libraryjournal.com/lj/home/889748-264/two_university_press_ebook_initiatives.html.csp



Mollie
M. Pharo



Marcia
Learned Au

E-books and the Public Library: the Evansville Vanderburgh Public Library Experience

MOLLIE M. PHARO AND MARCIA LEARNED AU

The Evansville Vanderburgh Public Library (EVPL) is a county-wide library district in Indiana serving nearly 180,000 residents (2010 US Census) from eight locations: a Central Library and seven branch libraries located throughout the community. The libraries are community gathering places with comfortable, inviting spaces for people to learn, meet others, explore the world, and connect to information, resources, and technology. The Library's mission is to promote and support: reading, lifelong learning, economic vitality, and cultural initiatives.

EVPL has been recognized nationally for its quality, fiscal responsibility, and exceptional service to the community. In addition to providing physical buildings and resources, EVPL provides electronic resources and services through its internet presence as well as via social networking sites.

History of e-books at EVPL

When Stephen King's *Riding the Bullet* was published in March, 2000, in e-book-only format, the Library downloaded free e-book reading software and copies of *Riding the Bullet* to several public computers on which customers could read the title. At the time we were not able to collect statistics on the number of people who actually read it sitting in front of one of those clunky terminals but to the best of our collective memory it was very few, if any. Staff opined and we agreed that it was not a format that people were ready to embrace.

Despite the lack of interest in the King title and still believing that e-books were ever nearer on the horizon, the Library purchased a RocketBook Reader later in 2000. The Reader was then used for demonstrations and also circulated among the staff so that they would become familiar with what we were sure would kickstart the run on e-books. Some staff members also downloaded e-book reading software such as Glassbooks Reader (now Adobe Reader) and Microsoft's e-book reader to their computers and became familiar with them. Unfortunately neither of these formats nor the RocketBook Reader ever took off—pun intended.

In 2001, the Library joined netLibrary via the INCOLSA network. This network included several other Indiana public libraries but because of the number of libraries participating, the ability to access the material became problematic. Bibliographic records for the 378 e-books in that collection were loaded into EVPL's catalog in June 2001. Beginning July

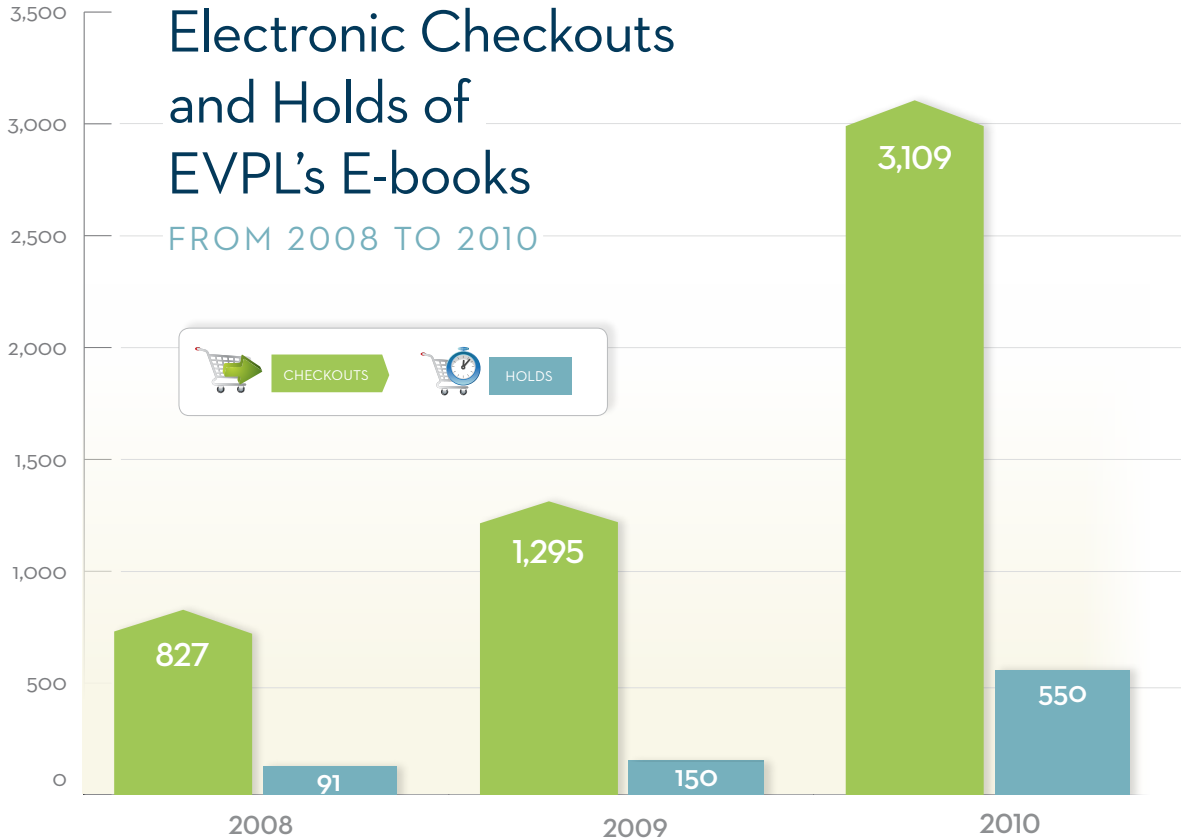


Figure 1: Electronic checkouts and holds

2001, EVPL customers were able to access netLibrary books held in the EVPL collection along with public domain e-books in netLibrary's collection. Problems arose when it became obvious that the demand created by multiple libraries sharing access to 378 e-books could not possibly be met nor were holds (reserves) available. Although Indiana's INSPIRE project now provides a statewide collection of NetLibrary e-books, EVPL is not currently adding netLibrary titles to our collection.

In 2005, the Library joined OverDrive along with three other libraries (via the INCOLSA network) using the shared collection model. In October 2006, EVPL left the shared collection when once again the number of titles could not meet demand and there seemed to be some confusion about collection development, i.e., which organization was responsible for selecting which titles. The realization that a shared collection was not working particularly well drove the Library to negotiate an independent contract with OverDrive. During the next four years, usage was predominantly of e-audiobooks rather than textual e-books. In an effort to push staff familiarity with technology mid-management personnel

were provided with Palm PDAs and, because of the format recommended for those and other hand-held devices and cell phones, e-books in the Mobi-Pocket format were purchased. The external interest in e-books, however, was still minimal and because of the relatively low collection use we seriously considered ending our relationship with OverDrive several times over those years. It was also during this time that reality struck; we realized how limited our options were—the content we'd considered as purchased from OverDrive could not be kept if we left them. Their e-content really was and is more of a leased product than a purchased one.

We made the decision to remain with OverDrive and that was advantageous for us as we built the e-book collection during those lean-use years. The greatest advantage was that we achieved a comparatively large selection of titles. When the tipping point came at the end of 2010 when e-book readers were the "de rigueur" gift for the holidays, the Library was able to supply more than just a basic collection and we had staff who were familiar enough

CONTINUED »

with e-books to do some hands-on assistance and provide e-reader advisory as well.

E-books today at EVPL

What did all those 2010 holiday e-reader gifts mean to the Library? Much as the publishing industry saw e-book sales jump 164.4% last year, EVPL e-book use (downloads from OverDrive) increased over 500% from January 2010 to January 2011, with EPUB format e-book downloads increasing over 800% from EVPL's OverDrive site in the last quarter of the year alone (see Figure 1).

To help keep staff abreast of the technology and developments with e-reader technology, EVPL continues to make investments in e-reader hardware. Currently the Library owns a Kindle, 2 Nooks, several Sony eReaders, a Kobo, two iPads, and a Pandigital eReader as part of our Tech Toolbox. Toolbox items are available for staff to check

out as well as use in programs for both staff and customers. The Library has been designated as one of 30 libraries accepted for the second round of Sony's Library Program. Since December 2010, a total of 15 e-reader and OverDrive download workshops for staff, customers, and other area librarians have been conducted. Staff continue to add more sessions on a regular basis. To date over 400 people have attended the workshops. Additionally Help Screens and handouts specific to the different e-reader devices have been created and widely distributed. E-books and audiobooks are featured on the front page of the Library's website. Customer comments indicate that they view library staff as unbiased experts on e-readers and staff have been inundated with questions about using OverDrive as well as how to operate the various e-reader devices. While staff do not promote specific e-reader products in workshops, they are able to talk about the differences in devices and which formats work

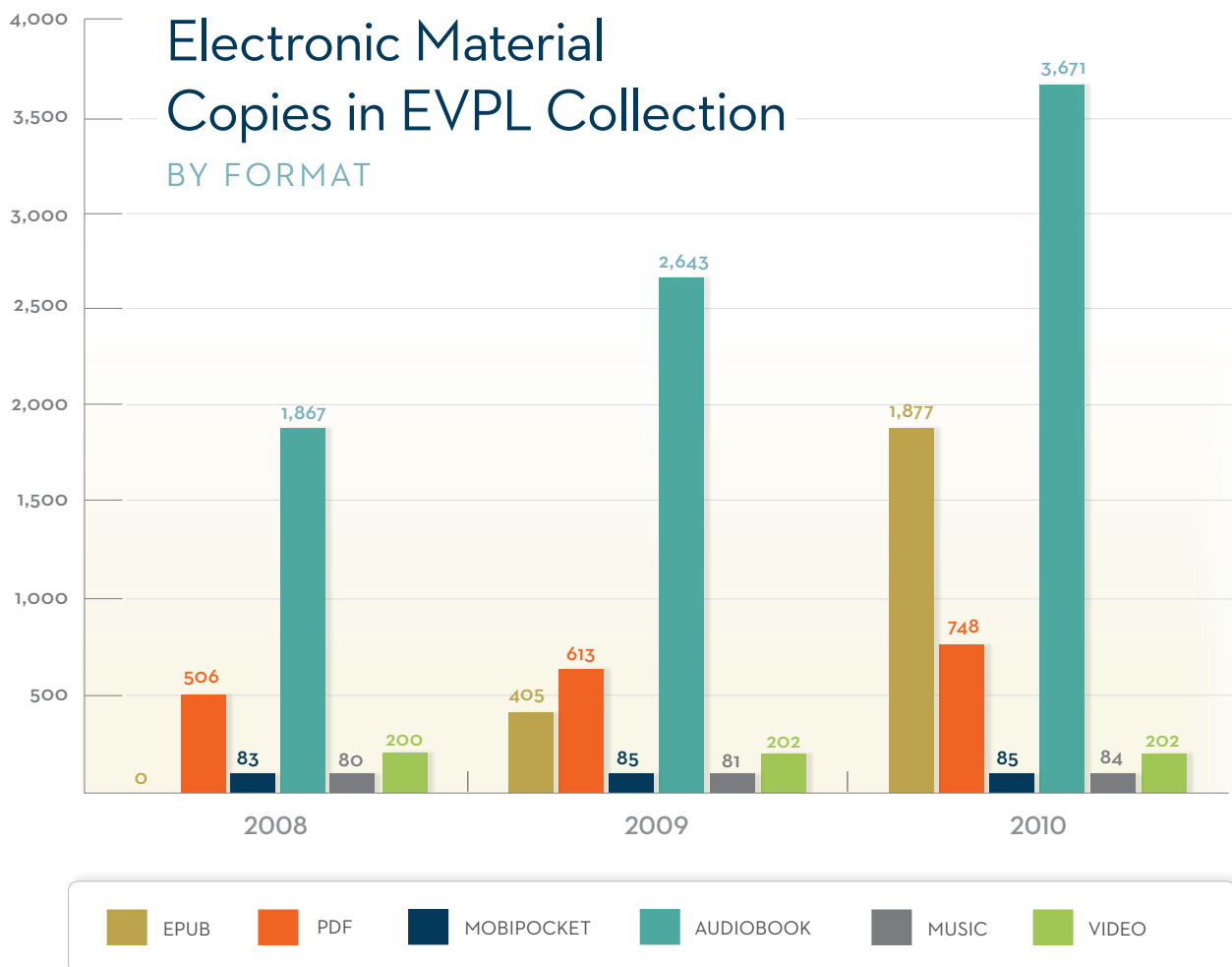


Figure 2: Electronic material copies in collection by format

with which e-reader. They also provide some basic education about what e-books are and realistic expectations about what customers will be able to download/borrow from the Library and what that experience will be like. We want customers to have a positive experience using the library's e-books, a challenge given some of the difficulties using the OverDrive platform and the fact that the Kindle is currently not compatible with content available to public libraries. EVPL also makes an effort to build working partnerships with our local book vendors, notably our local Barnes and Noble store. Barnes and Noble invites EVPL staff to the workshops they offer publicly that feature their e-reader (the Nook) and their e-book selections. In turn, Barnes & Noble staff participate in some of EVPL's e-reader workshops to talk about their product and how it works with library e-books. It is a win for both organizations, but an even bigger win for our customers.

EVPL has invested heavily in providing electronic content to the citizens of Vanderburgh County. We are convinced that if libraries can provide the electronic content that our customers want, when they want it, where they want it, and directly to the devices they want it on, the Library stands a much better chance of continuing to provide equitable access to these new resources and to play an important part in how individuals and businesses in our community are empowered and become successful.

What about standards?

One of the issues of ongoing concern is, "In what format should we purchase/lease e-books?" For some time we were purchasing/leasing mostly PDF and Mobi-Pocket format e-books. Our customers and staff did not use the Mobi-Pocket format, so we ceased acquiring those. Now we purchase/lease mostly EPUB format (see Figure 2). We are building collections for the majority of our customers who want to download e-books to their personal e-reader or tablet or cell phone, and who want the popular fiction titles. The e-reader devices our customers are using seem to handle the EPUB format most easily. We also switched to EPUB because it was touted as the new "standard" format, and standards can be very helpful and make things more accessible over the long-term. That said however, academic libraries seem to be leaning more toward the PDF format of e-books and so we ponder from time to time if we should also be adding titles in that format, or in this brave new e-world if some new format will supersede both. EBSCO has indicated that they are working to acquire or convert the e-books they will offer on their EBSCOhost/NetLibrary product to the EPUB format and, since we are subscribers, our decision is to stay with EPUB for now.

On a related note we recently stopped archiving copies of the e-books we purchase for our Gale Virtual Reference Library

collection, largely because we have no idea how we could ever use those outside of the Gale platform because of their formatting.

What about the future of e-books and other electronic content at EVPL?

"May you live in interesting times" is the adage that springs to mind when contemplating the forces driving technology today, resource access issues, and the mission of the public library. This is indeed a very interesting time in the evolution of electronic content, with e-books recently becoming more popular and in demand, and with publishers and authors moving to protect their streams of revenue as the interest in and use of e-books grows. We in public libraries find ourselves in a difficult position—apparently not seen as an important part of the revenue stream but instead often seen as siphoning off revenue by providing free loans to customers. As a result, or perhaps just as a result of not considering libraries, most models for electronic content delivery have been direct-to-consumer models. Two major publishers currently do not allow their e-book product to be purchased/leased to libraries and another has begun leasing their e-books with a total download/loan limit. Libraries are starting to work together, and with publishers and vendors, to provide "library-friendly" models to meet our customers' needs. The Library Renewal project, the Urban Library Council, the Chief Officers of State Libraries COSLA: *ebook Feasibility Study Final Report*, the Internet Archive's In-Library Lending Program, and the ALA's Presidential Taskforce on Equitable Access to Electronic Content (EQUACC), are all examples of this very important work. Vendors are also working on these issues, and it looks as though OverDrive will soon have some real competition in providing e-books and e-audiobooks to public libraries and library customers.

While demand for e-book titles continues to ramp up, expenditures on e-collections are growing more slowly. One reason for this is the economic concern that looms before us as a publicly funded entity. Do we own this electronic content or have we only leased the right to access it on the vendor's platform? What happens if we choose to leave that platform, migrate to another vendor, or develop an infrastructure that supports working with authors directly?

Digital rights management (DRM) requirements imposed by publishers complicate what vendors and libraries can provide. Customers would like to see DRM go away. Some authors are doing this, making their titles available for DRM-free download from their websites. Many authors are self-publishing their e-books, bypassing traditional publishers. Smashwords assists independent authors and small

CONTINUED »



amazon.com

KINDLE LIBRARY LENDING PROGRAM

Since submitting this article, Amazon and OverDrive have announced the Kindle Library Lending program, expected to enable e-book content in library OverDrive collections to be downloadable to Kindles later this year. Who knows what other game-changing announcements will be made before publication?

publishers with this effort. Small, independent presses and authors themselves seem more receptive to providing DRM-free content to libraries. At present, we view EPUB as our preferred e-book format, anticipating that DRM will not go away anytime soon for the library lending market.

Some libraries are now developing their own infrastructure and negotiating directly with publishers and authors for electronic content. Of note is the partnership being developed by the Colorado Independent Publishers Association (CIPA), the Red Rocks Community College, and the Douglas County Libraries. According to reports, by June 2011 the two libraries will offer e-books from CIPA's authors for checkout through their catalogs but will also allow click-through purchase of these titles. While not as formally developed as the Colorado model, EVPL is also investigating options for providing content from local authors through our Library catalog.

EVPL continues to monitor e-book related developments very closely. In many respects e-book development is still in its infancy. In the short term, technophiles anticipate that there will be significant changes in the hardware and some of us believe firmly in convergence theory, that in the very near future there will not be a plethora of e-book devices but one device that will be an e-reader, a telecommunications device, a wallet, etc.—but that is another topic for another article. In the meantime, we are determined to stay relevant, to continue the Library's mission, to address the legal and ethical issues of copyright, and most importantly provide access to resources for all people. We are staying tuned as they say and looking forward to the next round of developments in the brave new e-world.

IP | doi: 10.3789/isqv23n2.2011.04

MOLLIE M. PHARO <molliep@evpl.org> is Collection Development Manager and **MARCIA LEARNED AU** <mau@evpl.org> is CEO/Director of the Evansville Vanderburgh Public Library.

ALA Presidential Taskforce. Equitable Access to Electronic Content (EQUACC).
www.equacc.ala.org

Chief Officers of State Libraries (COSLA). eBook Feasibility Study Final Report.
www.cosla.org/documents/COSLA2270_Report_Final1.pdf

Colorado Independent Publishers Association e-Book Partnerships
www.cipacatalog.com/pages/Library-eBook-Partnerships.html

Evansville Vanderburgh Public Library
www.evpl.org

INSPIRE (Indiana's Virtual Library)
www.in.gov/library/inspire/faq.html

Internet Archive's In-Library Lending Program
blog.archive.org/2011/02/22/in-library-ebook-lending-program-launched

Library Renewal project
libraryrenewal.org

netLibrary
www.netlibrary.com

OverDrive
www.overdrive.com

Sony Library Program
ebookstore.sony.com/library-program/

Smashwords
www.smashwords.com

Urban Library Council
www.urbanlibraries.org



RELEVANT
LINKS

OP

[OPINION]

A judgement formed about something;
a personal view, attitude, or appraisal



Wendy Allen
Shelburne



WENDY ALLEN SHELBURNE

Drinking the E-book Kool-Aid in a Large Academic Library

I would love nothing more than to be able to say that I absolutely never met an e-book I didn't like because for the most part I simply love e-books. What really can be better than a book that an unlimited number of people can read at the same time from multiple locations, search and retrieve every word, and download or print as much as they need or want to? The book is never checked out or lost, there's more than likely no pages missing or odd notes from the previous user, and this book can be used whether the library is open or closed.

That's not to say that I think there isn't room left for improvement in e-books or that all of the e-book problems in the world have been solved, but for the most part even with their problems, let's face it, they are fantastic. When we line up our collective memories and experiences to consider what electronic resources and associated tools have done for the expansion of content to users, the ways in which they have facilitated and changed teaching and learning, and how they have affected the work of libraries and publishing, it is simply an amazing journey. And to this day, after eight years as an Electronic Resources Librarian, I am still waiting to meet the user who is not happy to be able to use electronic content. Given this, I am sometimes more

than a bit shocked, knowing that the "new" iteration of e-books is surely five to six years old at this point, that we are not further along on this shared journey, and we still haven't run out of things to say and/or do, or whine about with regard to e-books. But then I remember, we haven't run out of things to say about other electronic resources either, but they may not take up as much of the spotlight anymore.

The First Adventure, or Let's Buy a Big Package of E-books and See What Happens

At the University of Illinois I have been, for the most part, an evangelist for e-books since late 2005, when I was first

introduced to a new publisher-based e-books model soon to be on the market. This was the Springer e-book model. It was so very exciting to me as it held so much in common with the e-journal model and, knowing a thing or two about e-journals and e-journal usage, if nothing else I was pretty sure our users would be very accepting of an e-book that looked and smelled like an e-journal. Turns out I was not wrong, and our journey into large scale, frontlist e-book purchasing began. I should be truthful and admit, however, that no matter how exciting all of this was on the surface, it was incredibly difficult on the back end and, as a matter of fact, in some ways it was just downright ugly and remains so even today. Beginning with arguments about

CONTINUED »

how to realign budgets to pay for them, to how to make the duplicate print copies stop coming, to utter confusion about what titles were being published and which ones were we getting, and ending with a pretty sad tale about having no idea where to get the MARC records for our shiny new resources until many months after the fact. To be honest, it was really a complete and utter mess, at times almost verging on disaster, but in the end we stayed the course because everyone believed it was really worth doing no matter if we killed each other in the process. Today I would say our coping skills are significantly improved, but we still have a long way to go.

This Model vs. That Model vs. That Other Model

So, what is so hard about e-books outside of money? I am leaving money out as an obstacle to e-books since money is an obstacle to just about everything in libraries these days and it seems very disingenuous to single it out against e-books. First and foremost, the hardest thing about e-books in my mind is that there may be as many models to purchase and or subscribe to them as there are Dalmatians. Keeping them all straight is almost a feat unto itself. There are one-time purchase models from publishers, one-time purchase models from aggregators, e-books you can download and print just like e-journals, e-books you can only cut and paste from or only print one page at a time from, e-books that are HTMLfiles only, and e-books that are only

for an individual and not an institution. Or there are the e-books that you buy via a combination pricing model of users and title value that you can then weed later. Notice how I'm avoiding the device conversation, not to mention all of the other models I can't possibly list here. But the worst e-book model in all of this, in my opinion anyway, is the dreaded site licensed PDF download book. Imagine the standard workflow, as if e-resources actually had workflows, where staff purchased PDF e-books online and then we e-mailed them around to each other to load on web servers, add to proxy, and then hand-add links to A-Z lists, and, I assume, then add a catalog record as well as a record and target parser in the link resolver. The only thing worse that I can think of is the "free," for however long, PDF e-books online that are found and sent in to have us add them to the OPAC.

Currently I would guesstimate that I have at least one weekly interesting e-book "model" conversation since part of my job is to investigate resource pricing for our selectors. The one that is clearest in my mind at this writing is the one where the publisher has a print plus online model whereby you purchase the print copies for X dollars and have online access free for two years. I was very interested to find out what would happen at the end of the two years, but unfortunately the publisher did not know, and there was no option to just purchase the electronic. Yes, I did laugh out loud and yes, I did apologize for it, but seriously? Let me say it again, I want e-books that act like e-journals, and I

want to buy them like e-journals, even if it is just picking out one now and then as opposed to a package. And I want to pay my invoice and have someone load a record and be done with it so that I can move along to the next problem that is waiting to sneak up behind me.

MARC, Metadata, and Why Isn't It in the OPAC Yet?

After models, I suspect the next challenge with e-books is MARC records and managing discovery. From who makes them to where do we get them and ending with the never-ending discussion about quality, it seems that no one, even if they tried, could come up with a harder thing to do relative to e-books. Given that there is now access to electronic abstracts for books, and frankly more metadata than we ever dreamed of relative to book publishing, are we really going to continue with the MARC record as the gold standard of access to an e-book? It is unfathomable to me that almost nine years after Roy Tennant declared "MARC Must Die" I am still trying to figure out why it is so hard to load the MARC records for the e-books, and am now also watching the growing e-book cataloging backlog that might ultimately match those cataloging backlogs of print, which never quite seem to completely go away. It is also patently clear to me that it is a good thing I decided not to pursue a career in cataloging, as I obviously would have been really bad at it. Further, I have to say that I remain very worried that link resolvers and associated electronic resource management knowledgebases and tools will never truly match what is needed to manage e-books because we remain forever tied to MARC, and still somehow lack the ability to rethink not only how we buy these materials but how we provide access to them. Seriously, did we not learn anything at all from e-journals?

I am still trying to figure out why it is so hard to load the MARC records for the e-books, and am now also watching the growing e-book cataloging backlog that might ultimately match those cataloging backlogs of print, which never quite seem to completely go away.



I like to think of devices as eight-track tape players, but an even more clever colleague called them disposable razors and I'm sure I like that better. I just hope that in our environment we are able to purchase formats that allow us to remain device agnostic.



The Demand and the Spending Continues to Grow

Regardless of what is confusing, challenging, or just downright awful about e-books, my mantra at the end of the day is "This is what my people want." I certainly cannot speak outside of my experiences at Illinois for those are the only ones I know as a librarian, but the user community here is more than happy to tell us at any given time that they want e-books, and even more e-resources while we are at it, and they want more of them as soon as possible. Discipline concerns are also appearing to fade. The last e-book title I ordered, at the author's request (a faculty member at University of Illinois at Urbana-Champaign) is: *The Art of Building in the Classical World: Vision, Craftsmanship, and Linear Perspective in Greek and Roman Architecture*, published by Cambridge University Press. And when our users aren't specifically asking for the content, all that is needed is a quick check of the usage numbers and we know we are on the right path. Just last week, in participating in "Snapshot: One Day in the Life of Illinois Libraries" the quick stat from our A-Z list for e-book accesses for a single day was 472, and a year ago was 118. Overall, measureable use of e-books for a day totaled to 750. Our COUNTER stats for e-book usage confirm the desire for this content continues to grow, and in some cases it already rivals the usage of some e-journal collections. In terms of overall spending, the way our budget is built does not support much in the way of comparison for format spending, e.g., electronic vs. print, but it is already clear that in the current fiscal year (2011) our e-book spending will easily approach \$1 million.

Is It Going to Get Worse Before It Gets Better?

While we already know so much about e-books, and have certainly purchased a good many of them to date, there's still so much remaining development, evolution, and just plain ole stuff to learn. Publisher-direct, "all in" packages may prove unsustainable for both parties. It could be the case that patron-driven acquisitions (PDA) become a requirement for both aggregator- and publisher-based purchase programs. Consortial PDA models may additionally provide even more interest in this model, as if there isn't enough already. This will certainly assist in the problem of "sharing" e-book content, but a centrally or individually funded patron driven model across institutions may not fit the bill for all parties. What will be the future role of the book jobber, especially given new developments in ownership and development of aggregated platforms, e.g., the impending University Press E-book competition between JSTOR and the University Press Content Consortium (UPCC), as well as the inevitable clash of the vendor titans between ProQuest with Ebrary and EBSCO with NetLibrary. We may well see competition for e-book provision that was never witnessed in the e-journal environment. Finally, much as I would like to avoid the topic, the proof will be in the pudding when it comes to devices and the users will go where they go. I like to think of devices as eight-track tape players, but an even more clever colleague called them disposable razors and I'm sure I like that better. I just hope that in our environment we are able to purchase formats that allow us to remain device agnostic.

CONTINUED »



The current options for libraries to obtain e-book content run the gamut and we will likely never be able to say we did it exactly right, but we can at least say we did it to the best of our abilities given our resources.

CONTINUED »

Why It's Worth the Effort

I would like to think that for any library the ultimate goal is getting the users what they want. Unfortunately, our resources don't always match the need, but this has been the case for many years and our ability to keep pace with the body of publishing has always been challenged. There's certainly no silver bullet for curing what challenges libraries when it comes to e-books. Nonetheless, the current options for libraries to obtain e-book content run the gamut and we will likely never be able to say we did it exactly right, but we can at least say we did it to the best of our abilities given our resources. All of this can and does work, but it is just a mess in between. At Illinois we are pretty much working under every model we can make work, while ignoring providers with models we don't like, and waiting to see what opportunity will come around the bend. We buy books from large and small publishers in "all in" packages both solo and consortially; we are setting up a wide variety of standing orders; we are leasing content on subscription models; we are picking up one-by-one titles where that makes sense; and, like so many others, we also have a patron driven program. None of these could be considered perfect in any way. We struggle to keep up with the output of the larger publishers and are always behind in providing access when new titles are loading every day. We struggle to help our selectors understand what we have purchased and what we haven't purchased, even though we know full well something we didn't anticipate will be excluded for a reason we may not necessarily understand and we will likely have to buy it later. We wonder if we will ever have a mechanism for interlibrary loan, as we have received those rights in so many cases. We wonder what it will be like two years from now. We wonder what it will be like 20 minutes from now.

In the end, for all that can be said regarding e-books, probably the most unfortunate comment I ever hear is that "we don't have any," or "we don't want any." E-books can only be ignored at our own peril. There's simply too much at stake to not engage wholeheartedly in shaping our collective future, and e-books are just not going to go away this time.

I OP | doi: 10.3789/isqv23n2.2011.05

WENDY ALLEN SHELBURNE <shelburn@illinois.edu> is Electronic Resources Librarian, Acting Head of Acquisitions, and Associate Professor of Library Administration at University of Illinois at Urbana-Champaign <www.illinois.edu>.

Snapshot: One Day in the Life of Illinois Libraries

www.ila.org/snapshot/

Tennant, Roy. MARC Must Die. *Library Journal*, 10/15/2002.

www.libraryjournal.com/article/CA250046.html

Springer eBooks

www.springer.com/librarians/e-content/ebooks?SGWID=0-40791-0-0-0

JSTOR

www.jstor.org/

University Press Content Consortium

muse.jhu.edu/about/new/ebook_collections.html



RELEVANT
LINKS



Mark Bide

STANDARD SPOTLIGHT:
ISBN AND ONIX FOR BOOKS

MARK BIDE

The Challenge for Standards in the E-book Supply Chain

*The e-book supply chain is complicated—and is unlikely to get simpler any time soon. What do I mean by “the supply chain”? I mean the whole process that gets an e-book from author to reader—the only two **really** important points in the whole chain. Without authors who willingly write and readers who willingly read, there would be no supply chain to worry about. But our primary focus in this article is those intermediaries who add value in the process from author to reader. Ask any author who has stood on a street corner trying to sell (or even to give away) copies of a manuscript to passers-by. The process which gets a book from author to readers adds value.*

In the physical world, the supply chain involves a number of different types of intermediaries coming together in different combinations: literary agents, publishers, typesetters and book manufacturers, distributors, wholesalers, retailers, and library suppliers. Now add the digital distribution chain alongside it, with its range of service providers and aggregators (sometimes the same companies, sometimes completely different ones). The process of shaping and re-shaping the digital supply chain is far from complete in 2011—but at least in the short term it isn't getting any simpler.

In both supply chains, there are myriad organizations that need to be able to talk to each other, to exchange information about the stuff that passes through the supply chain—in other words, to exchange metadata. And this is where standards come into the picture: they provide the common language that allows us to speak across organizational boundaries from machine to machine, oiling the wheels of this vastly complex enterprise, ensuring unambiguous communication and (to the extent possible) friction-free commerce.

All too often, discussions of metadata focus on a single application—discovery. Of course, in the digital world, discovery is in some ways a greater challenge than it used to be in the physical one because the only tool you have to bring readers to authors is metadata. Thus all online merchandising and marketing is about the metadata. So publishers are increasingly taking all aspects of their metadata seriously. While discovery is a particular driver, we cannot forget that high quality, accurate metadata lies at the core of all automated business processes. And efficient and highly automated processes lie at the heart of successful commerce on the network.

At the core of the metadata that drives those processes lies identity...

CONTINUED »



ISBN has quietly created the backbone of all of our standards and all our systems in the book trade.

The challenge of identity

The ISBN—which is arguably the most successful product identifier ever devised—was introduced to replace individual publisher’s catalog numbers and to enable the first drive to electronic commerce. Without clarity of identity, it was not possible to use computers to manage the supply chain. The entire structure of EDI (electronic data interchange) standards on which an effective book supply chain has been built over the past 40 years has only been possible because of the implementation of the ISBN—distinguishing hardback from paperback, third edition from fourth edition.

ISBN has quietly created the backbone of all of our standards and all our systems in the book trade. This has been both our strength and our Achilles heel. The ISBN has enabled us to be sure that we are all “talking about the same thing,” but its utility has been such that we have used it for all sorts of purposes for which it was never designed. The ISBN was (and is) intended to identify products in the supply chain. Look inside most publishers’ systems, and you will find the ISBN used as a proxy to identify all sorts of things that are not products. It is not atypical, for example, for publishers to have a control on their cost ledgers such that it is not possible to incur costs on a publishing project without an ISBN. Whilst it may not need saying, a project and a product are simply not the same thing.

None of this mattered unduly when there was typically a close correlation between the “content” of the product and the “product” itself, and when the variety of products of any given “title” that could be made available was limited—maybe a hardback, a trade paperback and a regular paperback. The hardback ISBN was often used as the “master ISBN,” to collocate (aggregate information about) this limited number of products (for example, in royalty ledgers). But note the sudden rash of quote marks in this paragraph. We are beginning to move into areas of uncertainty—areas where the meanings of words become uncertain and, critically, often mean different things to different people.

This is the type of ambiguity which is extremely threatening to the efficient operation of e-commerce systems. Computer systems are not good at resolving ambiguity and uncertainty. While we were dealing with physical products, the impact of this ambiguity was reasonably well controlled and rarely surfaced as a problem outside the walls of an individual publishing house (where it was rarely recognized for what it was). With the advent of the “e-book,” however, the problem is suddenly becoming rather more acute.

The e-book and the ISBN

The answer to the question, “How do we identify our e-books?” seems very obvious. Use the ISBN. But it turns out it isn’t quite as simple as that.

The first issue is the lack of clarity of what distinguishes one e-book from another e-book—at the product level. When the ISBN standard was last revised, e-books were still nascent. Although the current edition was published in 2005, the primary work on revising the text inevitably predates the formal publication (as anyone who has ever been involved in the creation or revision of an ISO standard will well understand). At that point, differences between e-book products were seen as analogous to the differences between a hardback and paperback—and the distinction that is drawn in the standard is between different technical file formats (the examples including PDF and HTML, as well as a number of file formats now obsolescent or obsolete). Perhaps understandably, what could not be foreseen at that time was that the development of the e-book market would not entirely mimic that of the physical book market, and that critical differences in the supply chain would make the application of a different ISBN

to each product much more challenging than it might have appeared in those early days.

There are several contributory factors to the rather unsatisfactory position in which we now find ourselves as an industry. Why unsatisfactory? Because what has previously been a reasonably consistently implemented standard, has evolved to a situation where there are widely differing practices in terms of ISBN allocation in different markets, between different publishers in the same market—and even sometimes between different parts of the same publishing house. Because of these differences in policy and practice, we are losing the certainty of identity that the ISBN normally affords us.

This is the first major upset that I can recall for the ISBN since the fierce debates with designers in the early '70s about the damage wrought to the artistic integrity of cover designs by putting barcodes onto books. This may be hard to imagine now, but it was very real then—at least until (in the UK at least) a dominant retailer announced “no barcode, no sale”; this closed the argument down very effectively. There are perhaps lessons to be learned from that experience.

Why is there a problem with e-book identification?

There are at least three major challenges with identifying e-book products with ISBNs.

The first is relatively straightforward and has already been mentioned. How do you distinguish one e-book product from another e-book product? The answer to this question has been provided in a set of Guidelines published by the International ISBN Agency. Although these guidelines may need to be further extended and nuanced over time, it appears that the general concepts that underlie them are proving robust.

However, although there is growing consensus at the theoretical level, there are still serious barriers to implementation. The first (and perhaps the most difficult) is that publishers do not know a priori exactly what products will be created from any given content, and cannot therefore easily pre-allocate ISBNs to the different products at an early point in the production lifecycle. ISBNs rather need to be available “on the fly” when the requirement for an additional product is identified. Unfortunately the creation and delivery of e-book products are typically not undertaken by the publisher, but by a digital service provider working on the publisher’s behalf, or by an aggregator. There are no mechanisms available to the publisher—or to the service provider—to facilitate the issuing of these ISBNs at the appropriate point in the lifecycle (in other words, precisely when they are needed). One solution to this has been to allow these intermediaries to have their own prefix and to apply ISBNs themselves to publishers’ products. But despite some successful implementations of this model

(for example, by O’Reilly in their Safari online book product), it remains generally an unpopular option, particularly with publishers, not least because of the problems it creates for management of product metadata records. (It is not unusual for publisher systems to be unable to manage ISBNs issued by other publishers.)

Which takes us to the second problem that proliferation of products implies: this is frequently (pejoratively) referred to as “metadata bloat”—as if, somehow, metadata itself is growing out of control, a malign presence in the basement of the industry. Of course, the problem is that if you have a more complex world and more complex business, your metadata simply reflects that complexity. You don’t simplify something by simplifying its description; with that approach you simply lose knowledge (data) about whatever it is you are describing, and this sort of data, once lost, is often impossible to regain.

This is not to suggest that there isn’t a real problem here. Systems designed to manage a simpler world are often not appropriate for managing the sort of complexity that we are now facing. Many publishers’ systems create metadata records for a new product by “cloning” the record of a related product and then editing the fields that identify the differences between the products. In the case of different e-book products, these may be very small differences. But now, instead of small numbers of metadata records for “the same” title, you have a growing number of individual records. And any time a change has to be made, there is no way of editing these records as a batch; each has to be individually edited, which is not only time consuming (and therefore expensive), but also error prone.

Although system solutions to this are in the development/deployment pipeline for the major vendors of publishing systems, it will be a while before they are anywhere near universally deployed. Quite apart from anything else, there is limited appetite for investment in systems at a time of considerable uncertainty. It is understandable that when the e-book market is doubling or tripling in size annually, grabbing market share and managing that growth takes precedence over any efficiency there may be in the better management of metadata. And, of course, it has to be recognized that the current explosion of different non-interoperable e-book products may be a passing phenomenon, with some sort of convergence point in the middle distance—in which case, why expend effort on a passing phase?

So, while some publishers have continued to recognize the importance of managing their different lines of product by effective identification, we have seen others deploying a single ISBN for all e-books (sometimes dubbed an “eISBN”). The one thing that is for certain is that there is no such thing as an eISBN—even if people are using a 13-digit number that looks like an ISBN—because this identifier doesn’t identify a product (the only class of entity that an ISBN can be used to identify) but

CONTINUED »

rather a class of products sharing the same content but different product attributes. Unfortunately, the idea of the eISBN has even reached the world of MARC cataloging (where its use has been promoted, despite the fact that it doesn't exist!).

And it isn't that these two positions (one ISBN for each e-book product, or one for all e-book products) are the only two models being followed. The reality is much more complex than this, with almost every imaginable practice being followed (including some publishers who persist in identifying their e-books with the hardback ISBN).

So, what we are faced with is an industry that has slowly seen its primary identifier system—once the flagship of identifier standards—slide into a chaos of incompatible practices and “workarounds.” While I remain optimistic that we will establish international agreement on the implementation of ISBN, cleaning up the aftermath of the inconsistencies of the last four or five years is a different matter.

ISTC—an answer to the problem?

A contribution to resolving the challenge of collocation in the e-book market—drawing together the multiplicity of products containing the same content—could lie with broader application of the International Standard Text Code. This standard identifier, from the same ISO Committee which looks after the ISBN, is for identifying textual works—the abstract “content” rather than any specific manifestation of that content in a particular product.

However, this is another standard that is finding it hard to generate significant market traction. There are numerous possible explanations for this, but on the basis of recent research with publishers the requirement for a standard work identifier seems pressing. This apparent mismatch is deserving of further exploration.

E-books and ONIX

Having thought about some of the identification challenges we face, I will turn to the wider metadata picture. And for EDItEUR that means ONIX, and specifically ONIX for Books. The roots of its development lie in the 1990s, with the recognition by the Association of American Publishers (AAP) that there was a growing need for publishers to be able to communicate “rich product metadata” to online booksellers in an XML messaging format. The first release of ONIX was developed by the AAP, and the standard was then passed for long term governance to EDItEUR where we have managed it ever since. ONIX for Books 2.1 (released in 2004) has been widely deployed around the world; it is a credit to its designers that recent deployment in Japan has required minimal amendment to the standard.

However, it became clear about four years ago that ONIX required a major upgrade—and although in ONIX for Books 3.0



we have attended to a number of other weaknesses identified in earlier versions, the major driver behind the upgrade was the need to improve the capability to describe e-books. We undertook a major overhaul of the standard, with the approval and indeed encouragement of our International Steering Committee that represents the very large number of ONIX for Books implementers worldwide. However, release 3.0 does represent a significant challenge because in order to achieve the requirements identified by our users, we deliberately chose an upgrade structure that was not backwards compatible.

This creates a challenge for early adopters: why implement a messaging standard which no one can yet receive from me or send to me? Particularly when in ONIX 2.1 I can do 80% or more of everything I want to do...

This situation helps to explain the much slower progress towards release 3.0 implementation than hoped for when we launched it two years ago. We are now beginning to see more widespread implementation with the first major trade publishers following the pioneers in the academic market. I am optimistic that this will reduce the numbers of times that people ask me: “Why can't we describe this in ONIX?” when what they mean is “Why can't we describe this in ONIX for Books 2.1?” Because then I can stop giving the slightly frustrated answer, “Because you haven't implemented ONIX 3.0.”

However, one challenge will keep recurring unless we create a significant change; ONIX for Books manages records at the product level. And here we return to ISBN country. We have recently once again been asked, “How do you describe more than one product in a single ONIX for Books record?”—a question to which the answer is (for the questioner) frustratingly clear: “You cannot.” ONIX for Books is (and always has been) about product description—and a product (by definition) can only have one set of ONIX descriptors. A group of non-interoperable e-book products, with variations in file format, technical protection, usage limitations, hardware, or software requirements, and so on, remains a group of products and must be described with a group of product metadata records, notwithstanding that they manifest the same content.

About two years ago, recognizing and understanding the requirement for description of product groups, we put



There is considerable divergence in ONIX messaging practice, even within, for example, English language publishing. Messages from a US publisher cannot be interpreted by a recipient in the same way as messages from a UK publisher. This was fine while markets were organized nationally, but is posing an increasing challenge as the market and many of the key players in it become global.

a proposal for solving this dilemma to the ONIX for Books National Groups—the organizations represented on the ONIX for Books International Steering Committee (our governance group)—and it was unanimously rejected.

So we continue to face something of a quandary. Our pragmatic driver in standards development is to meet our stakeholder requirements; but our constraint is that our stakeholders find a degree of consensus. Part of our role is to facilitate that consensus—but that can be difficult when attitudes have become so polarized.

Growing pains

It is perhaps inevitable that the fundamental changes to the book industry that the “switch to digital” represents will be accompanied by some apparent lack of coherence when seen from the point of view of an organization whose role is to provide standards support. We are a long way from an understanding of how—or indeed whether—the shape of the market will settle down. However, one thing is becoming increasingly clear: markets are becoming increasingly global.

ONIX for Books has always been organized on the assumption that local implementations would vary from country to country and that “best practice” guides would be created at a national level. As a result, there is considerable divergence in ONIX messaging practice, even within, for example, English language publishing. Messages from a US publisher cannot be interpreted by a recipient in the same way as messages from a UK publisher. This was fine while markets were organized nationally, but is posing an increasing challenge as the market and many of the key players in it become global.

Our response has been to launch the first ever set of international best practice guidelines for ONIX for Books. While these will undoubtedly need to be supplemented locally—particularly in the physical book supply chain, where local practices will continue to need to be supported—we are optimistic that we can begin to resolve some divergences which have not been driven by any real differences in requirements, but simply by habit. The switch to ONIX for Books 3.0 is a real opportunity to improve consistency.

This is an essential step towards achieving another of our targets: more effective compliance. It is another commonly heard complaint that, “No one implements ONIX in the same way.”

There are at least three possible explanations for this:

- 1 Inadequate or imprecise documentation, either from EDItEUR or from national groups
- 2 Imperfect implementation, based on developers not following documentation—either “guessing” at what things mean or because of a need to work around system inadequacies
- 3 Demands of powerful individual players in the market for customized data feeds, which are difficult to resist by smaller organizations (or even larger ones anxious to get their products to market) but which lead to a fracturing of the standard

To the extent that the first of these is in our own hands, we are doing what we can to improve documentation through the publication of the international best practice guidelines.

We can also help to some extent with the second of these challenges by giving direct support for implementations to our members (something we are offering on an increasing basis) and by publishing improved compliance testing tools—exemplified by our work on a Schematron schema for ONIX for Books 3.0, which enables users to validate messages against a much wider variety of parameters than either a DTD or simple XML schema.

The last challenge is more difficult. Ultimately, compliance is a peer-community challenge more than it is a central “enforcement” one. We cannot act as policemen; we can only exhort all those who implement our standards to be more forceful in driving out non-standard implementations—otherwise, the cost savings available through the implementation of standards can never be optimized.

CONTINUED »

Major recipients of ONIX data have a critical role to play here in encouraging data providers to adhere to broadly accepted standards—and to best practice—rather than demanding idiosyncratic proprietary interpretations. And we are keen to help these major recipients in their interpretation of the standards under our care, so our stakeholders can avoid costly recipient-specific metadata.

Meeting the challenge of convergence

The issue of the requirement for convergence arises throughout this article. In an increasingly global market, we need convergence between different organizations, and between different countries, in the way they implement the same standard (whether we are talking about ISBN or ONIX for Books). But convergence is going even further.

We are seeing convergence between requirements for ONIX for Books and those for what have traditionally been called ONIX for Serials messages. This family of messages—designed for communication within the library supply chain—was always exclusively focused on journal subscription products. However, we have recently undertaken a substantial overhaul of these messages to allow for them to cover any type of content that is provided on a subscription basis—including e-books and databases—and indeed non-textual resources. Although we have no present intention of enriching the ONIX for Serials messages with the sort of detailed product information that can be communicated in ONIX for Books, the message of the market is clear: the tidy distinctions between books and journals are rapidly being broken down. (There is also nascent interest in using ONIX to communicate about subscription products in the consumer market.)

And this brings us to the two final points that I want to make about convergence. The first is between ONIX for Books and MARC. These two standards have developed in very different ways—for good reason. There is a marked difference between requirements for book marketing and requirements for book cataloging and the different standards reflect these. Nevertheless, there is a dawning recognition of the potential for closer collaboration “across the divide.” The work that OCLC has done in developing the ONIX to MARC (and back again) crosswalks is symbolic of this, as is the Library of Congress use of ONIX to improve the efficiency of its CIP program. EDItEUR is a partner in a European project called Linked Heritage which started in April 2011; our role in this project is to find ways to bridge the gap between commercial metadata and the Europeana digital library. All perhaps slightly tentative first steps, but all pointing in the same direction.

The final convergence challenge is perhaps the most significant but at the same time even more challenging to address than the differences between ONIX and MARC: convergence between the different media. Now that they can all be “consumed” on the same electronic device, it is proving

increasingly difficult to draw the clear distinctions that we once so easily made between different media types. As the channels to market converge, it is entirely unrealistic to believe that we can continue to ignore the challenge that standards convergence will pose for us. We are only at the very beginning of this process and the journey in front of us remains obscure; but it is a journey on which we need to embark sooner rather than later.

A simpler life?

I cannot see any real likelihood that things are going to get radically simpler in the immediate future. Nevertheless, I remain optimistic that the challenges that we are facing—complexity, compliance, and convergence—are all actively on the agenda. EDItEUR is working in ever closer collaboration with its members and with other standards organizations all around the world and in all the different media to find ways to resolve our common challenges. The next few years will continue to be very active ones in the standards community.

ISPI doi: 10.3789/isqv23n2.2011.06

MARK BIDE <mark@editeur.org> was appointed Executive Director of EDItEUR in January 2009; he remains a Director of Rightscom, the specialist media consultancy where he has worked since 2001. He is a Visiting Professor of the University of the Arts London.

EDItEUR

www.editeur.org

Guidelines for the assignment of ISBNs to e-books

isbn-international.org/faqs/view/17

International ISBN Agency

isbn-international.org/

International ISTC Agency

istc-international.org

Linked Heritage

www.cyi.ac.cy/node/1094

Mapping ONIX to MARC [OCLC]

www.oclc.org/research/publications/library/2010/2010-14.pdf (report)

www.oclc.org/research/publications/library/2010/2010-14a.xls (crosswalk)

ONIX and MARC21

www.editeur.org/96/ONIX-and-MARC21/

ONIX for Books

www.editeur.org/11/Books/

ONIX for Serials

www.editeur.org/17/Serials/

Provider-Neutral

E-Monograph Record (This July 2009 report refers to something called an eISBN, while making it clear it is not product-specific.)

www.loc.gov/catdir/pcc/bibco/PN-Final-Report.pdf

Safari Books Online

my.safaribooksonline.com/

Using ONIX with Cataloging in Publication (CIP)

cip.loc.gov/onixpro.html



RELEVANT
LINKS

MEMBER SPOTLIGHT:
EBSCO PUBLISHING



Michael Gorrell
Senior Vice President,
Chief Information Office,
EBSCO Publishing

E-books on EBSCOhost:[®] Combining *NetLibrary*[™] E-books with the EBSCOhost Platform

As libraries expand their services and keep up with technology, vendors that serve libraries need to keep up as well. E-books and audiobooks are becoming core services for libraries and as patrons and researchers turn to their libraries for these resources, libraries are in turn looking to their vendors to meet those needs. At EBSCO Publishing (EBSCO), we see e-books as an essential resource for libraries. All of the types of libraries we serve are buying e-books and audiobooks, which are rapidly growing parts of their collections.

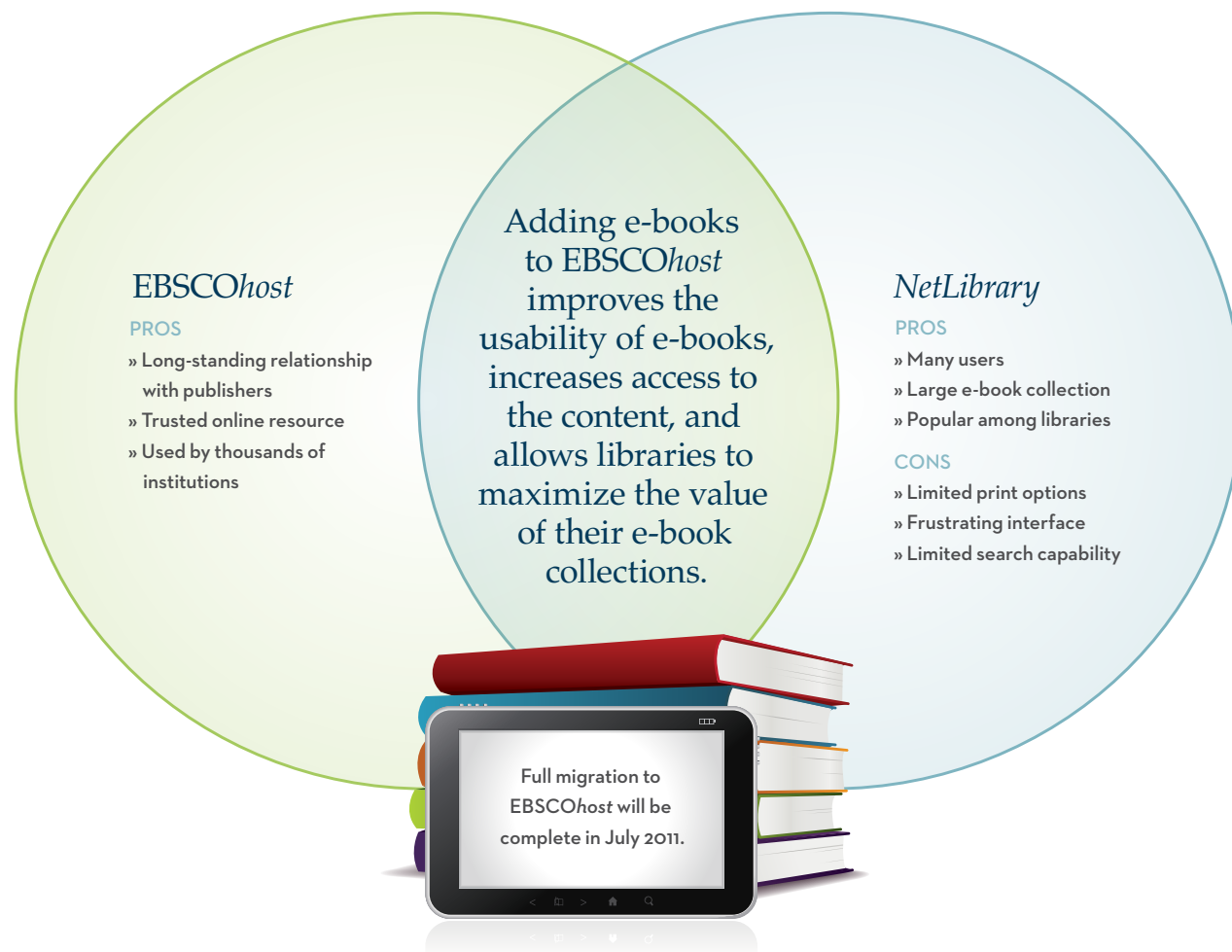
Given our long-term relationships with both publishers and libraries, we see ourselves as being well-positioned to help both parties by leveraging the value and popularity of the EBSCOhost[®] platform as a delivery vehicle for e-books. Publishers benefit from having an aggregator that understands their needs and libraries benefit from having a vendor and a platform they know and trust to provide a variety of e-book options that allow libraries to make the best access decisions to serve their end users.

The actual idea of adding e-books to EBSCOhost began with customer suggestions that we become an e-book aggregator. Once we established that the idea had merit we began looking at the landscape. Acquiring *NetLibrary*[™] in March 2010 provided us with a unique opportunity: take the original e-book platform to new heights with an extended EBSCOhost platform that will provide the best e-book and audiobook experience for libraries. It also gave us the chance to leverage our great relationships with publishers to extend the substantial content base that came with *NetLibrary* in addition to greatly expanding the licensing models under which content is available. At this time EBSCO was also busy with the continued development of *EBSCO Discovery Service*[™] and as we looked ahead, we realized that both initiatives helped libraries benefit from the improved integration of content (electronic and print) through a unified interface that offers the end user ease of access along with the sophistication to provide a familiar, powerful user experience.

When EBSCO Publishing acquired *NetLibrary* we had the long-term goal of integrating e-books into EBSCOhost while at the same time making immediate improvements to the *NetLibrary* platform. An example of improvements that were

CONTINUED »

Acquiring *NetLibrary*[™] in March 2010 provided us with a unique opportunity: take the original e-book platform to new heights with an extended EBSCOhost platform that will provide the best e-book and audiobook experience for libraries.



made included expanding print options from 15 pages up to 60 pages. At the same time, we were even more focused on analyzing and revamping all of the *NetLibrary* systems and processes to make them the most modern and optimized ones in the industry. All of this development was designed to achieve our ultimate goal—enabling users to search e-books and audiobooks alongside their EBSCOhost databases.

The preview of e-books on EBSCOhost, which was released in March 2011, was a year in the making. The debut of e-books on EBSCOhost incorporates suggestions from librarians all over the world and includes the results of user testing designed to create a seamless and integrated environment that supported searching databases as well as e-books. The preview also introduced the new e-book viewer designed to maximize the user's experience with each e-book. While *NetLibrary* is one of the best known e-book platforms for libraries, the *NetLibrary* interface has been a well-known source of frustration for librarians and end users. *NetLibrary* was simply in need of updating to introduce better searching, faster speeds, and more options for librarians and end users

alike. EBSCOhost was primed to extend its position as the most-used research platform.

While *NetLibrary* has many users, traditionally e-books have been searched separately on their own interface. EBSCO's goal was to integrate the content into EBSCOhost, which is used in tens of thousands of libraries, and to combine the large e-book collection available from *NetLibrary* with the powerful search experience available from EBSCOhost. Adding e-books to EBSCOhost improves the usability of e-books, increases access to the content, and allows libraries to maximize the value of their e-book collections. The most significant improvement that librarians and users will experience now that e-books are available on EBSCOhost will be the functionality that the EBSCOhost platform brings to e-books including such resources as the ability to save, e-mail, and view citations and bookmark results. Users will also be able to explore the table of contents from the result list or the detailed record more easily. (In *NetLibrary* browsing the table of contents required the user to access the full text, which meant the title was unavailable for other users.) Another

new feature is the inclusion of BISAC categories, which introduces new user-friendly subject headings and search functionality and takes library classification to a new level.

Once the full migration to EBSCOhost is complete in July 2011, we will introduce an optional e-book landing page and an optional audiobook landing page to facilitate content browsing. The site will also feature full 508c accessibility, COUNTER-compliant reporting, permalinks to page numbers, and branding options. Later in 2011, enhancements will include the ability to load EPUB content, which will rapidly expand the number of front-list titles available on EBSCOhost and will strengthen the downloadable e-book program. Additional enhancements not related to the interface will include expanded access models such as unlimited user models, multiple user models, short-term lease models, and improved patron driven acquisition (PDA) options.

Since access models depend on publisher agreements, EBSCO is leveraging its long-standing relationships with publishers to negotiate a variety of access options for libraries. The current *NetLibrary* standard of one book, one user is limiting. New options include a three-user model and an unlimited-user model. EBSCO will also offer upgrade options. A library could start out with a three-user model and upgrade to an unlimited model when librarians see that the demand warrants expanded access for a title and want to retain ownership of the title while also offering it to each user without delay. In much the same way, a *smart* PDA approach would allow a site to establish which titles its collection development experts want to expose to PDA and allow patron usage to drive both acquisitions and upgrades. A library might initially pay for one user and establish a framework for upgrades in which a certain number of subsequent uses or requests would trigger an upgrade to a three-user or an unlimited-user model.

Collection development for e-books and audiobooks on EBSCOhost can be achieved in a variety of ways. Currently the TitleSelect service allows libraries to search nearly 300,000 titles available from EBSCO and make selections on a title-by-title basis. Subject Sets are a prepackaged set of titles chosen specifically for their subject appeal. The EBSCO collection development team of librarians and collection specialists use their expertise and knowledge to create collections and Subject Sets for libraries. To date, EBSCO has created more than 150 e-book Subject Sets—representing more than 2,500 titles. All Subject Sets from EBSCO include titles published within the past three years and have no duplication among current or past offerings. EBSCO also creates Standard Collections which work as great starting points for a library to begin selecting a wider range of titles within a given discipline. We offer more than 200 Standard Collections, each of which has between 500 and 2000 titles. Lastly, we offer Custom Collections, which are built by EBSCO's collection development librarians based on the distinct needs of a given institution. This approach is designed for larger purchases and lets libraries work with our collection development librarians, sharing their content objectives and knowledge of their current collections along with their budget in order to create a custom collection for the library to consider. The idea is that the library then works with this list of suggested titles to either purchase in whole or refine. In the end, this is still "title-by-title" acquisition as the library can elect to remove or add any titles it sees fit.

After the migration of *NetLibrary* to EBSCOhost, a new EBSCOhost Collection Manager (ECM) will be introduced to take the place of TitleSelect. EBSCOhost Collection Manager will provide a simpler user experience for librarians. ECM will make it even easier for librarians to create or add to their collections on their own,



Collection development for e-books and audiobooks on EBSCOhost can be achieved in a variety of ways. Currently the TitleSelect service allows libraries to search nearly 300,000 titles available from EBSCO and make selections on a title-by-title basis. Subject Sets are a prepackaged set of titles chosen specifically for their subject appeal.

CONTINUED »



E-books on EBSCOhost are designed to meet the needs of any library and its users...As EBSCO grows its e-book and audiobook collections, we will continue to add titles that serve the needs of academic, medical, and corporate users but the addition of EPUB titles will allow EBSCO to add more of the popular fiction and non-fiction titles that public libraries require.

build or augment collections with Subject Sets, create a profile and be alerted when new titles or collections meeting the profile become available, participate in collaborative collection development with colleagues, or work with EBSCO to create a custom collection.

E-books on EBSCOhost are designed to meet the needs of any library and its users—from serious academic researchers looking for the latest information on a subject who will now see e-books integrated with periodical content to public library patrons looking for an e-book or audiobook to download to the latest mobile device. Currently, Adobe®'s digital publishing solution facilitates the download for e-books on EBSCOhost allowing access from most dedicated e-book readers and mobile devices, including Sony eReader, Nook™, Samsung eReader, iPad, iPhone, and Droid phones. As EBSCO grows its e-book and audiobook collections, we will continue to add titles that serve the needs of academic, medical, and corporate users but the addition of EPUB titles will allow EBSCO to add more of the popular fiction and non-fiction titles that public libraries require.

Libraries are in many different phases of adding e-books into their collections and they want a variety of options from their e-book provider. Some libraries want to fully integrate e-books and databases, while others want separate profiles. EBSCO has always offered strong administrative capabilities through EBSCOadmin™ and we will offer a variety of options for e-books. We believe we are offering librarians the best interface to bring these resources to their end users along with the most varied access models. We will continue to work with libraries and publishers to improve access and the options that are available to help libraries add e-books and serve the needs of their end users. After having so many librarians suggest that we begin to provide e-books on EBSCOhost, we are eager to hear what our library customers think and our goal is to continue to develop an e-book service that provides the best user experience and unparalleled content along with the most flexible acquisition and licensing arrangements.

ISPI doi: 10.3789/isqv23n2.2011.07

MICHAEL GORRELL <mdg@ebSCOhost.com> is Senior Vice President, Chief Information Office, EBSCO Publishing.

EBSCOhost

www.ebscohost.com/

EBSCO Discovery Service™

www.ebscohost.com/discovery

Net Library

www.netlibrary.com



**RELEVANT
LINKS**


 Matt
Garrish

 Markus
Gylling

MATT GARRISH AND MARKUS GYLLING

The Evolution of Accessible Publishing: Revising the Z39.86 DAISY Standard

As you read through the revision of the DAISY standard, *Authoring and Interchange Framework Specification* (NISO Z39.86-201x) and its profiles, rest assured it's no coincidence that the markup examples are drawn predominantly from works of Charles Darwin. This specification has undergone a radical transformation since the Working Group began its work in the Fall of 2008, and the result represents a significant evolutionary leap forward in accessible content production.

Darwin's *On the Origin of Species* was consequently selected as the primary source for examples as a quiet nod to the principles of adaptation and evolution that this specification has embodied over the years. This new revision represents a whole new way of looking at the parallel-publishing model in particular, and at content model creation in general, but wouldn't have been possible if not for its predecessors on the road to universal accessible publishing.

of the specification quickly became the de facto standard for talking book production by libraries and organizations serving blind, dyslexic, and other print-disabled readers.

Although a very effective and time-tried specification (it remains in use by many producers to this day) and the one that ushered in the age of digital talking books, the community creating and using these books also had a need

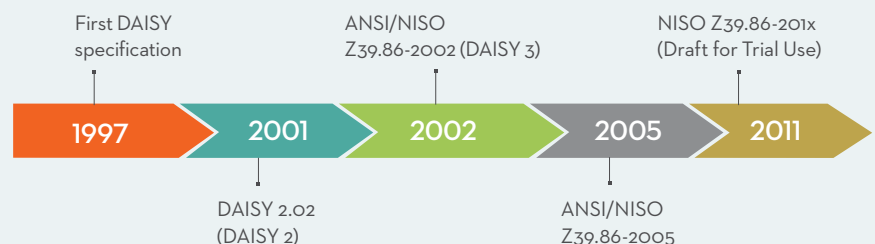
to generate other formats from their text data. This task of reformatting documents was often a repetitive one that involved a combination of machine and human intervention. Producers were increasingly looking to their DAISY text files as the source for these conversions, to leverage the cost and effort that had already gone into converting the original documents they represent into digital form. But while HTML is fine for the Web, it didn't take

CONTINUED »

A brief history of Z39.86

Accessible format production has come a long way since the first *DAISY Digital Talking Book (DTB)* specification was developed in 1997. That early format combined the HTML and SMIL (Synchronized Multimedia Integration Language) standards to create a synchronized multimedia experience that was ahead of its time, and after a few early revisions the 2.02 version

TIMELINE OF THE DAISY STANDARD





While the markup that it provided proved efficient in many authoring contexts, it was insufficient in others; the requirements for formats like braille and large print were still not adequately addressed for all producers.

CONTINUED »

long before it became clear that a more structured format with better facilities for targeting outputs was going to be needed to enable multi-format production. (Attempting to generate print braille using the small tag set HTML makes available can prove no small feat, for example.)

From this need was born the ANSI/NISO Z39.86-2002 standard and its subsequent 2005 revision. The new text component of these versions of the standard was defined by the DTBook DTD, which built on the original HTML core but added significant improvements for structural and semantic fidelity. As DTBook was deployed by producers of accessible content across the globe, it clearly showed how producers could benefit from XML-based single-source production, and how end users benefit from textual content that is well-structured and semantically coherent.

The wind of change

But although DTBook again improved the production landscape, it brought forward with it the specification's legacy of talking book production and the Web. While the markup that it provided proved efficient in many authoring contexts, it was insufficient in others; the requirements for formats like braille and large print were still not adequately addressed for all producers. Its *book-centrism* and limited mechanisms for adaption and specialization additionally meant that it was inadequate to handle all the document types and regional requirements of producers.

Meanwhile, in the end user context, DTBook as a distribution format anticipated browser vendors moving to accommodate display and rendering of arbitrary XML grammars. This shift—which was seen as a given a decade ago when DTBook was originally created—never materialized, proving a major complication for the visual rendering of talking books. Further, the DAISY distribution format (as of Z39.86-2005) lacked several highly requested features such as

interactivity and better support for East Asian languages.

By 2008, it had become clear that the text component had to be thoroughly revised and cleanly separated from the talking book format if it was going to meet the multi-format production needs that the community was clamoring for, and the distribution format needed to be re-aligned with industry standards.

Evolution in action

The first decision made in undertaking the revision of the 2005 standard was to adopt the principle of *separation of concerns*: to split the incongruous parts in order to isolate and better tackle the problem domains. A new XML authoring standard would be developed to address the accessible text production needs of the community, while a distribution format—a more linear continuation of the previous standard—would focus on talking book production.

The next critical decision in designing the new text standard was that the past would not be a guide to the future. A radical departure was instead needed if the DAISY Revision Working Group was going to be successful. To this end, it was decided that creating a specific markup grammar was not going to be the primary goal of the revision.

While this might seem like a strange objective for a text standard, the group had to invert the production problem and look at it from a fresh perspective. The single monolithic format approach had so far failed to address the needs of the broad community DAISY serves, providing neither the structural and semantic richness nor the flexibility to accommodate the wide array of formats producers had to be able to generate. To begin developing yet another such standard would be to head down an evolutionary dead end.

To fully realize the benefits of a parallel publishing model, a true master source was needed that provided a content model that wasn't hampered by the formatting inherent in being embedded in a specific output or being designed for

a single use. It was envisaged that this new specification would define a common framework in which to develop new grammars, allowing the standard to be adapted to any document type it had to address, instead of the other way around. The framework would specify the technologies to use and define a universal markup core for all documents, but would stay silent about how to structure any given type of document: the structure would be left to be defined by profiles created according to the rules laid out in the framework.

This approach would provide the increased flexibility that producers were requesting to allow markup to be tailored to their unique needs. Understanding that the community shared the same core markup requirements and that the incompatible differences largely related to output production requirements, a common framework was seen as a means of allowing producers to work collaboratively on profiles that fit their shared needs, or to strike out on their own but in a manner that still kept their core markup in line with the wider community. This consistency was going to be key to adoption in a community moving in the direction of a global library, where knowing the differences in markup between any two profiles, producers could still easily exchange and transform their documents to account for the discrepancies.

But this model is intended to be useful beyond just accessible publishing. Too often, the only solution when marking up new document types is to either: a) create a whole new markup model from the ground up; or b) find the closest fitting language and hack a solution over top of it. The new Z39.86 model encourages new profiles to be developed by any interested parties for the benefit of the whole community, sharing knowledge and enhancing the existing foundation to expedite the process. Although initially targeted at the accessible publishing community, the markup is designed to capture the full structure and semantics of the information resources being described, allowing any formats to be generated from them. Adoption of Z39.86 beyond its traditional base is key to making publishing open to all, and it is hoped that all organizations with similar cooperative markup needs and goals will benefit from the work put into this specification.

The next goal of this revision was to move the DAISY standard away from the legacy DTD approach that had persisted from its earlier HTML days. The XML ecosystem has largely outgrown DTDs, and their lack of native namespace and datatype support alone made them an incompatible choice for the direction the group was heading. W3C XML schemas were also discounted as the right technology for defining the lexical constraints on markup models. Although more progressive than DTDs, their unique particle attribution limitations in a document context (where like-named elements in block and phrase contexts are common) were deemed to be too limiting to make them a viable choice.

The Data Schema Definition Languages (DSDL) framework was instead turned to as a model for the future. Combining RelaxNG schemas for the structured markup together with Schematron assertions for enforcing finer markup logic provided exactly the balance of power and flexibility that was going to be needed for the modular framework that was planned.

Next generation of markup

Knowing how the group was going to implement the standard still left a long road ahead to build it. A model framework had to be constructed, rules for creating profiles defined, and working implementations developed that proved the framework was more than just an elaborate theory. The DAISY Revision Working



The single monolithic format approach had so far failed to address the needs of the broad community DAISY serves, providing neither the structural and semantic richness nor the flexibility to accommodate the wide array of formats producers had to be able to generate. To begin developing yet another such standard would be to head down an evolutionary dead end.

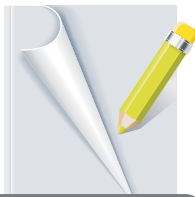
CONTINUED »



book profile



periodicals profile



document profile

These profiles were targeted at the most prominent information resource types the community handles: a book profile for general works of fiction and non-fiction, a periodicals profile for news and magazine articles, and a general document type for word processing and similar documents found in office environments.

Group spent the next two and a half years filling in these blanks.

The Abstract Document Model, the theoretical model underpinning the specification, was developed to define the basic requirements all profiles had to adhere to. This model defines the common structure that all Z39.86-compliant profiles must implement (i.e., the root element and metadata and body content containers). Existing document definitions were analyzed in developing this model, and from this research a common layering of structural elements became apparent: sectioning, block, phrase, and text. These layers were then formalized into the model to ensure that markup is always structurally consistent across profiles.

To facilitate the modular, plug-in architecture of the framework, a set of core modules was also developed to accompany the specification (i.e., the set of pre-defined components that could be drawn on when building new profiles, reducing the work involved in creating profiles and ensuring greater consistency between them). These components allow the rapid development of new profiles, as they can be included in any new markup model and tailored to the needs of the resource being described without having to be completely rewritten.

RDF (Resource Description Framework) metadata was also given a prominent place in the new specification. All profiles must include a minimal set of RDF support for header metadata, and hooks into the document structure through metadata attributes are also provided. An RDF profile must be defined for each markup profile, which contributes to the consistency of prefix naming across documents and simplifies implementation for document creators. The Working Group also undertook to create an extensive structural vocabulary of properties to augment the markup with additional semantic meaning (one that can address both mainstream and accessible publishing needs).

To prove that this model could work for real-world production, a catalog of profiles was developed in parallel with the specification using the technologies and rules outlined in it. These profiles were targeted at the most prominent information resource types the community handles: a book profile for general works of fiction and non-fiction, a periodicals profile for news and magazine articles, and a general document type for word processing and similar documents found in office environments.

The profiles were built using the same common module pool, but the content models they define are uniquely crafted to the resources they define—proof that this new approach was working as designed. Sample documents were likewise created using these profiles to ensure that the content models were rich enough to support real production. After much review by the working group members and organizations and the release of three public working drafts, the profiles have now been made available for test use by the community as part of the current review of the specification to gain additional feedback.

Building a better DTB through EPUB

Having discussed text at length, the question so far left unanswered is what happened to the digital talking book portion of the specification. Originally envisioned as a Part B distribution format, work on this specification was suspended as it became apparent that the new EPUB 3 revision was open to incorporating even more of the essential functionality of DTBs, with the goal of turning it into a specification accommodating readers of all abilities.

Rather than create a competing specification, principals in the DAISY Consortium began working in earnest with the International Digital Publishing

Forum (IDPF) to pool their resources to forge a joint standard, one that could address the cross-cutting requirements of both constituencies: a single e-book format that meets the needs of all readers (no more delays producing accessible versions), and a single recognized accessible format that publishers can produce and distribute, thereby reducing their costs.

No longer a mixture of e-book and DTB technologies, this new revision of EPUB has seen the DTB accessibility components more fully integrated into the specification:

- » The Navigation Control Center for XML Applications (NCX)—the menuing system for talking books—has been reformulated as an XHTML document to simplify its processing and rendering and to improve its international language capabilities.
- » The subset of SMIL used for synchronization of audio and text content, now called a Media Overlay document, lives outside the content markup. The provision of audio and text synchronization has generated substantial interest from mainstream publishers, proving that it is not a feature of interest only to print-disabled users.
- » Support for Text-to-Speech (TTS) markup has been integrated, allowing producers to enhance the content with pronunciation and prosody instructions.

The DAISY Consortium consequently anticipates adopting this new revision of EPUB as the distribution format for its members once the specification reaches recommendation status.

The End?

The Z39.86-201x *Authoring and Interchange Framework Specification* was recently released as a Draft Standard for Trial Use and a six month review is currently underway. The DAISY Revision Working Group anticipates being able to submit the specification for approval by NISO voting members and then ANSI after the trial closes on September 28.

While work will continue on the profile catalogs and core modules long after the specification itself becomes a standard, the anticipation is that this revision of the standard will provide a solid base on which the community can build their text production systems for many years to come.

The EPUB 3 family of specification documents are set to be released as formal recommendations during summer or early autumn of 2011, and their adoption by mainstream and accessible producers should be a swift process thereafter if early buzz is any indication. (See separate article on EPUB on page 4.)

But all things must evolve to stay relevant, especially standards... | NR | doi: 10.3789/isqv23n2.2011.08

MATT GARRISH <matt.garrish@bell.net> is an independent consultant who has done work for both the DAISY Consortium and the International Digital Publishing Forum (IDPF). **MARKUS GYLLING** <markus.gylling@gmail.com> is Chief Technology Officer at the DAISY Consortium.

DAISY Revision Working Group Members

Markus Gylling
DAISY, TPB (lead)

Josh Altherr
GH Braille

Ole Holst Andersen
Nota

Marisa DeMeglio
DAISY

Christian Egli
SBS

Matt Garrish
Invited expert

Boris Goldowsky
*Center for Applied
Special Technology*

Leona Holloway
Vision Australia

Kenny Johar
Vision Australia

Dennis Leas
GH Braille

Sam Ogami
*California State
University*

James Pritchett
*Learning Ally (formerly
RFB&D)*

Kathryn Randall
Vision Australia

Per Sennels
Huseby

Christian Waldfoegel
SBS

Z39.86 Authoring and Interchange Framework Specification

www.daisy.org/z3986/2011/Z3986-2011A.html

Z39.86 Profile Catalog

www.daisy.org/z3986/2011/auth/profiles/

Z39.86 Feature Catalog

www.daisy.org/z3986/2011/auth/features/

Z39.86 RDF Vocabulary Catalog

www.daisy.org/z3986/2011/vocab/

Z39.86 Core Modules

www.daisy.org/z3986/2011/auth/cm/

DAISY Consortium

www.daisy.org

International Digital Publishing Forum

www.idpf.org

EPUB 3.0

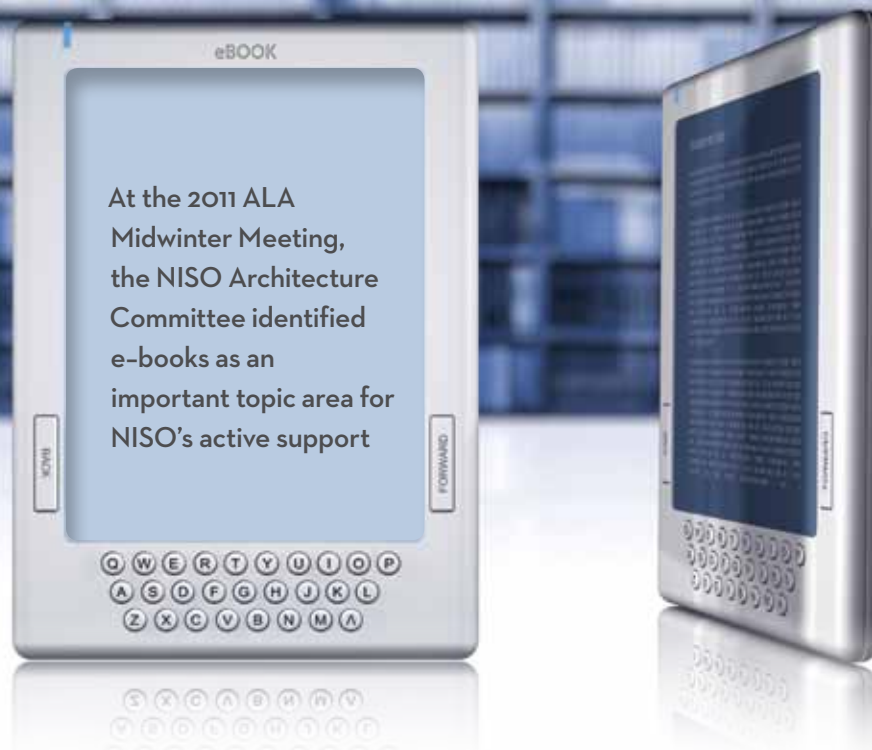
idpf.org/epub



**RELEVANT
LINKS**



Todd
Carpenter



TODD CARPENTER

NISO E-book Special Interest Group

While NISO is best known for shepherding groups focused on a specific problem through the creation and implementation of recommended practices and standards, NISO as part of its new strategic plan is seeking to widen its outreach within the information industry. Specifically, the NISO Board of Directors and Architecture Committee seek to emphasize NISO's role as a facilitator within the industry, one that can foster cross-community dialogue in a given topic area and provide a place for the incubation of ideas—even if no formal standards process within NISO is ever initiated as a result.

One outcome of the NISO Architecture Committee meeting at the 2011 ALA Midwinter Meeting was the identification of e-books as an important topic area for NISO's active support, even beyond its current activities. Another outcome was the acknowledgement that, although the current Topic Committee structure (Business Information, Content & Collection Management, Discover to Delivery) works well for most purposes, the investigation and exploration of more complex topics that cut across two or more of these groups (such as that of e-books) cannot be easily accommodated in the current structure.

For these reasons, NISO is creating an E-book Special Interest Group (E-book SIG) that cuts across all three Topic Committees and includes stakeholders from across the industry. The focus will be on both e-book content as well as delivery.

Currently, NISO is already engaged in e-books in a number of areas including NISO and ISO standards projects related to:

- » Formatting, markup, and distribution
- » Education
- » Licensing
- » Publications

The E-book SIG, which will help to coordinate and support these current activities, will also be exploring the larger context in which events are unfolding, including the following foundational questions:

- 1 How can NISO actively facilitate cross-community dialogue in the e-book area, building bridges between what are now separate, sometimes disparate groups?
- 2 How can NISO work collaboratively to provide education and information to assist with this dialogue?
- 3 How can NISO actively foster "incubation teams" to identify specific pain points in the e-books realm that could be remedied through formal standards, recommended practices, dissemination of information (e.g., via white papers, educational workshops, professional forums, Thought Leader meetings, etc.) either through NISO or another agency?

To start, the NISO E-books SIG will identify and start outreach to the communities within the library, publishing, vendor, and trade industries actively engaged in some aspect of e-books development and support. | NR | doi: 10.3789/isqv23n2.2011.09

We encourage any organization actively engaged in this area and wanting to be kept actively informed of the NISO E-books SIG's work to contact the NISO office at nisohq@niso.org.

E-BOOK SIG

Content + Delivery



NISO is creating an E-book Special Interest Group (E-book SIG) that cuts across all three Topic Committees and includes stakeholders from across the industry. The focus will be on both e-book content as well as delivery.



[NOTEWORTHY]

The JATS trial period will enable users to test the standard in real-life implementations and provide feedback.



All the schemas and tag libraries are in the public domain.

NLM Tag Suite 3 will be updated to version 1.0 of Z39.96 when approved.

There's a JATS Conference on September 26-27, 2011.

NISO ISSUES

Journal Article Tag Suite Standard for Trial Use

JATS: Journal Article Tag Suite, NISO Z39.96-201x, has been issued by NISO as a new draft standard for trial use through September 30, 2011. JATS provides a common XML format in which publishers and archives can exchange journal content by preserving the intellectual content of journals independent of the form in which that content was originally delivered. The draft standard defines elements and attributes for describing the textual and graphical content of journal articles as well as some non-article material, such as letters, editorials, and book and product reviews. The trial period will enable users to test the standard in real-life implementations and provide feedback. Once feedback is reviewed and NISO finalizes the standard, the final version will be submitted to ANSI for approval as an American National Standard.

The *Journal Article Tag Suite* standard is a continuation of the work started by the National Library of Medicine (NLM) on the *NLM Archiving and Interchange Tag Suite*, commonly referred to as the NLM DTDs. The NLM DTDs were based on an article model that was being used in the National Center for Biotechnology Information (NCBI)/NLM PubMed Central project to archive life science journals. The original PubMed Central article model was expanded in scope with support from Harvard University Libraries and The Andrew W. Mellon Foundation, in collaboration with Inera, Inc. and Mulberry Technologies, Inc., resulting in 2003 in the full *NLM Journal Archiving and Interchange Tag Suite*. The Tag Suite had reached version 3.0 prior to initiation of the NISO standardization

process; the NISO Z39.96-201x standard contains updates to version 3.0 of the NLM Tag Suite and is fully backwards compatible with it.

"As we looked at updating the NLM Tag Suite, we felt it was mature enough to take through NISO's standardization process," explained Jeffrey Beck, NCBI Technical Information Specialist at the National Library of Medicine and co-chair of the NISO JATS Working Group. "This standardization will bring awareness of the Tag Suite to a larger and more varied audience and provide opportunities for use in new applications."

"Since the release of version 1, the Archiving and Interchange Tag Suite has been widely popular," stated B. Tommie Usdin, President of Mulberry Technologies, Inc. and co-chair of the NISO JATS Working Group. "The format is being used to tag thousands of journals worldwide and is used for the journal archives at PubMed Central and Portico and by the online publisher HighWire Press. The Library of Congress and the British Library have announced their intention to use these models for archiving electronic content." ■

ⓧ The draft standard for trial use is available as both an online XML document and a downloadable PDF from the NISO website (www.niso.org/standards/z39-96/). An online commenting form is also available for trial users to provide feedback. Supporting documentation and schemas in DTD, RELAX NG, and W3C Schema formats are available at: jats.nlm.nih.gov/.

COUNTER Publishes Timetable and Objectives for Code of Practice Release 4

The fourth release of the COUNTER *Code of Practice* will combine the two existing codes for *Journals and Databases* and *Books and Reference Works*. The goal is to have a single, unified Code for all types of content, including multimedia works. The functionality of the XML reports and the use of SUSHI will be more fully developed and exploited. The implications of the PIRUS2 individual article usage statistics and the Journal Usage Factor projects will also be considered. The aim is to publish the definitive version of Release 4 in early 2012, with implementation by vendors required by December 31, 2013.

A timetable is provided to reach the goal of release in early 2012 and vendor implementation by December 2013. Both the existing Release 3 of the *Code of Practice for Journals and Databases* and the existing Release 1 of the *Code of Practice for Books and Reference Works* will remain valid until December 31 2013. ■

For more information on COUNTER, visit:
www.projectcounter.org/

For more information on PRIUS2, visit:
www.cranfieldlibrary.cranfield.ac.uk/pirus2/

For more information on the Journal
Usage Factor project, visit:
www.uksg.org/usagefactors/

The goal of this release is to have a single, unified Code for all types of content, including multimedia works.

RFID in Libraries – New Three-Part ISO Standard and NISO Revised Recommended Practice

The International Organization for Standardization (ISO) has published a new three-part standard, ISO 28560, *Information and documentation – RFID in libraries*, that specifies a model for the use of radio frequency identification (RFID) tags for library items designed to ensure interoperability between libraries in exchange of library items with RFID tags. *Part 1: Data elements and general guidelines for implementation* specifies the data model, system data elements, and user data elements to be used on the RFID tags. Two encoding methods are defined. *Part 2: Encoding of RFID data elements based on rules from ISO/IEC 15962* uses an object identifier structure to identify data elements. *Part 3: Fixed length encoding* deals with the encoding of a basic set of data elements in a fixed length format and the rest of the data elements in optional extension blocks. Parts 2 and 3 are mutually exclusive; the RFID tag would be encoded using only one of the two defined schemes.

To conform to the new ISO standard, NISO revised the 2008 recommended practice, *RFID in U.S. Libraries*, designed to be the U.S. profile for implementation of the ISO 28560 international standard. The international standard offers two different encoding options and many optional data elements; the NISO recommended practice recommends the use of the ISO 28560 Part 2 encoding scheme and further refines the usage of data elements so that U.S. implementers can adopt a common approach. In addition to the data model, the Recommended Practice provides guidelines on security, privacy, vandalism, and migrating existing library RFID implementations to the new model. Use of these recommendations will ensure that U.S. libraries can procure tags and equipment from different vendors, merge collections containing different manufacturers' tags, and, for the purposes of interlibrary loan, read the tags on items belonging to other libraries.

At the time of writing the NISO Recommended Practice (NISO RP-6-201x) was in the final week of a public comment period. It is expected to be published in June. ■

For more on the ISO standard, visit the ISO website (www.iso.org) and search on 28560.

For the NISO Recommended Practice, visit the working group webpage: www.niso.org/workrooms/rfid/

JSTOR FEATURES

15,000

New publishers, together with the original group, will result in more than 15,000 front and back list titles available.

1 million

Ability to search more than 1 million book reviews.

1,400

Citations dating back hundreds of years in the 1,400 journals available for JSTOR.

JSTOR Expands Scholarly E-book Offerings

Four more publishers announced plans to bring their scholarly books online at JSTOR. This is the second wave of presses to join the Books at JSTOR initiative. The initial group included Chicago, Minnesota, North Carolina, Princeton, and Yale University Presses.

JSTOR initially began archiving and bringing online the back issues of leading journals in economics and history. The focus is similar with books. Publishers are being invited to join the initiative based on the relevance of their titles to the content on JSTOR and importance of their publications to scholars now and in the future.

The new publishers, together with the original group, will result in more than 15,000 front and back list titles being made available at JSTOR.

These presses are recognized as leaders in a wide range of disciplines including:

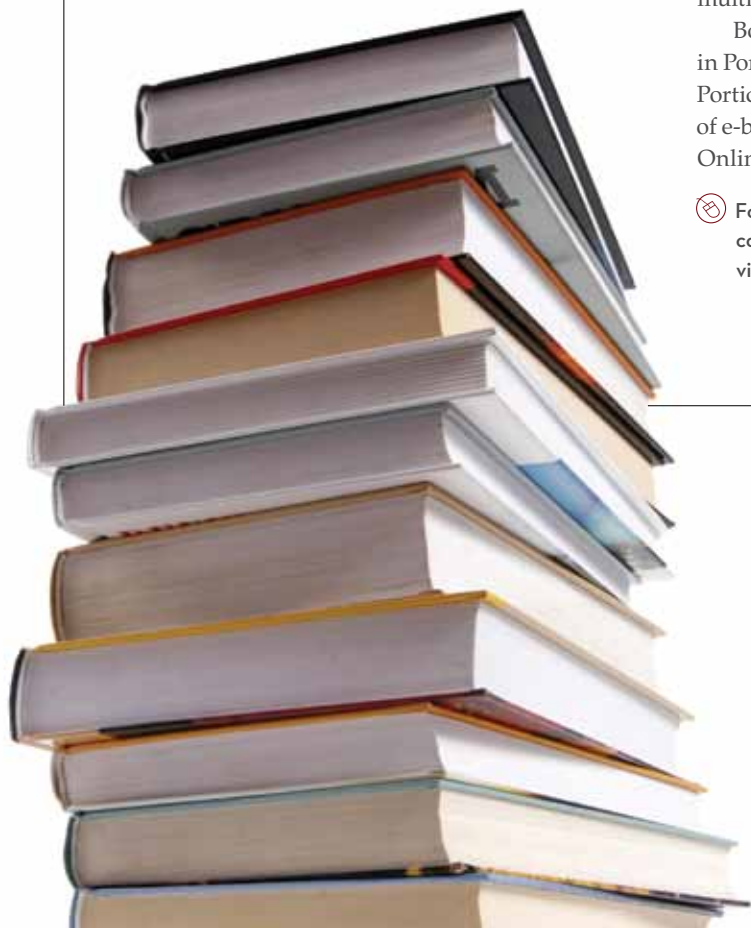
- » History
- » Sociology
- » The Arts
- » Business and Economics
- » Life sciences

In addition to offering the ability to search across journals and books, JSTOR includes more than 1 million book reviews and a vast number of citations dating back hundreds of years in the 1,400 journals it makes available. The books will be linked with this other content, creating a multitude of pathways through the literature.

Books will be available at JSTOR in spring 2012 and will be preserved in Portico, a digital preservation service for the scholarly community. Portico also announced that it will be preserving the entire collection of e-books from Oxford University Press from its Oxford Scholarship Online resource and Handbooks Online resource. ■

Ⓜ For more information on Books at JSTOR, visit: about.jstor.org/content-collections/books-jstor. For a list of e-book publishers participating in Portico, visit: www.portico.org/digital-preservation/who-participates-in-portico/

Four more publishers announced plans to bring their scholarly books online at JSTOR. This is the second wave of presses to join the Books at JSTOR initiative.



ProQuest Acquires ebrary

ProQuest through its acquisition of ebrary, a leader in providing e-books to libraries, will be adding over a quarter million e-books to their online content collection.

ProQuest plans continued investment in ebrary’s popular products and services for the academic, corporate, and public library markets including Academic Complete™ the company’s flagship product. ProQuest will also expand ebrary’s selection of research tools and ability to support new e-book devices as well as broadening language coverage from its current support of major European languages to include Chinese, Arabic, and others. The company will also accelerate the indexing of e-book content on its own new platform where books offered by ebrary will be searchable along with ProQuest’s existing research content of books, journals, dissertations, newspapers, video, government documents, and more.

Since the acquisition announcement, ebrary has continued to add new content and features, including: popular fiction e-books for the Public Library Complete product, a Topic of the Week program highlighting relevant e-books and offering free Title Preview™, new starter packs in 25 high-use subject areas, a patron-driven acquisition offering in conjunction with YBP Library Services, a usage-driven short-term loan model, and new German publishing partners along with a German-language interface. ■

🔗 For more information on ProQuest, visit: www.proquest.com
For more information on ebrary, visit: www.ebrary.com

Print Isn't Dead, says Bowker's Annual Book Production Report

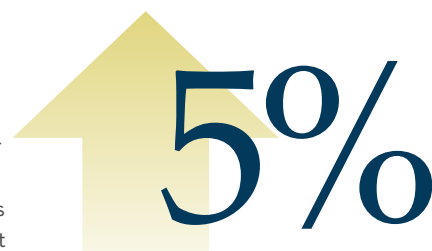
Bowker released its annual report on U.S. print book publishing, compiled from its Books In Print® database. Based on preliminary figures from U.S. publishers, Bowker is projecting that despite the popularity of e-books, traditional U.S. print title output in 2010 increased 5%. Output of new titles and editions increased from 302,410 in 2009 to a projected 316,480 in 2010. The 5% increase comes on the heels of a 4% increase the previous year based on the final 2008-2009 figures.

The non-traditional sector continues its explosive growth, increasing 169% from 1,033,065 in 2009 to an amazing 2,776,260 in 2010. These books, marketed almost exclusively on the web, are largely on-demand titles produced by reprint houses specializing in public

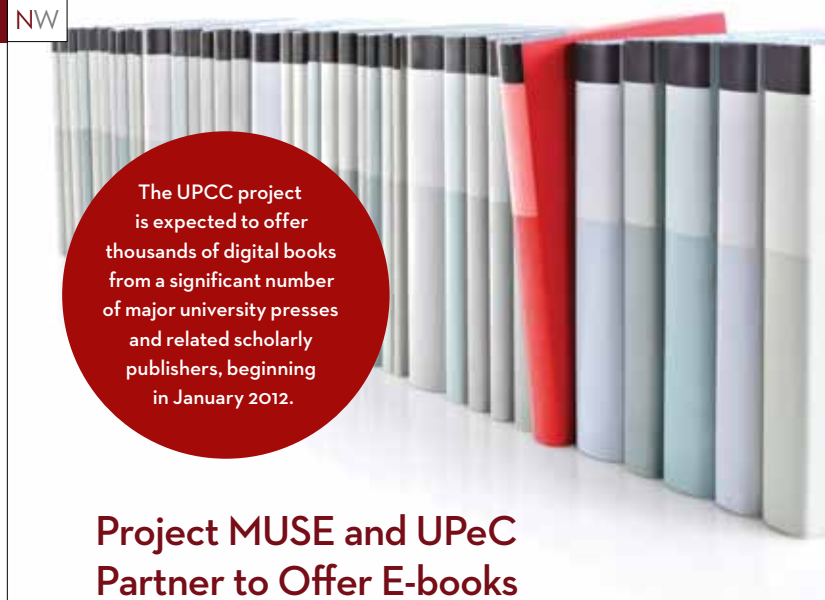
domain works and by presses catering to self-publishers and "micro-niche" publications. In 2008, the production of non-traditional print-on-demand books surpassed traditional book publishing for the first time and since then, its growth has been staggering. Now almost 8 times the output of traditional titles, the market is dominated by a handful of publishers.

In traditional publishing, SciTech continues to drive growth. Fiction, which is still the largest category (nearly 15% of the total) dropped 3% from 2009, continuing a decline from peak output in 2007. ■

🔗 For the complete report, visit: www.bowker.com/index.php/press-releases/633-print-isnt-dead-says-bowkers-annual-book-production-report



Based on preliminary figures from U.S. publishers, Bowker is projecting that despite the popularity of e-books, traditional U.S. print title output in 2010 increased 5%.



The UPCC project is expected to offer thousands of digital books from a significant number of major university presses and related scholarly publishers, beginning in January 2012.

Project MUSE and UPeC Partner to Offer E-books through the University Press Content Consortium

Project MUSE, an online platform initiated for humanities and social sciences e-journals, is further expanding into e-books with a newly formed partnership with the University Press e-Book Consortium (UPeC) to launch the University Press Content Consortium (UPCC). MUSE had previously announced their own Project MUSE Editions e-book program, but the new UPCC will replace that program.

The UPCC project is expected to offer thousands of digital books from a significant number of major university presses and related scholarly publishers, beginning in January 2012. E-books from 2010 and 2011 will be available immediately in January 2012 and archival backlists are also being developed. A beta platform including sample book content, and a new integrated search interface, will launch late this summer for review and feedback.

Books in the UPCC Collections will be fully integrated in the MUSE platform; users will be able to search across combined book and journal content, or limit searches by content type. Digital books will be in PDF format, searchable and retrievable to the chapter level, and will be released electronically at the same time the print book is released. Unlimited simultaneous usage of book content will be allowed, with no DRM and no restrictions on printing or downloading. Institutional purchasers of MUSE/UPCC book collections will receive ownership and perpetual access rights for books purchased. Free MARC records will be provided for all books and COUNTER-compliant usage statistics will be available. ■

For more information about UPCC and a list of publishers committed to provide e-books, visit: muse.jhu.edu/about/new/ebook_collections.html

Duke Releases Results of Library E-book Acquisition Survey

Duke University Press has published the results of an online survey regarding e-book acquisition by libraries. The survey link was sent to the Press librarian contacts, the LibLicense and ERIL listservs, and to their Twitter feed, resulting in 265 respondents.

Libraries are utilizing more than one acquisition strategy at the current time with 67% subscribing to aggregated collections from multiple publishers, 64% subscribing to single publisher collections, and 27% using patron-driven acquisition. Budgets are shifting somewhat: 53% indicated an increase in their budget for e-books. Many libraries are purchasing at least some portion of e-books in print format as well: 25% for highly-used titles; 19% for particular disciplines; 44% if requested by faculty or librarian; and 10% if the print versions are heavily discounted. The top three ranked criteria for selecting e-books are the quality of the content (87%), pricing (84%), and patron or faculty request (62%). The content platform at 55% and digital rights management (DRM) at 42% were the next highest reasons. In a separate question, 25% of respondents indicated that DRM restrictions were never acceptable. 70% of respondents do not provide patrons with e-book readers. Of those who do, Kindle and iPad were the top two e-readers offered.

Numerous write-in comments from respondents are included in the full survey report; they highlight the many issues that libraries are struggling with including pricing models, DRM, format and e-reader device variability, and the availability of cataloging records. ■

The survey report is available at: associationofsubscriptionagentsintermediaries.cmail5.com/t/r/l/jrlyjtt/ntykkgm/jj/

Cloud-Sourcing Research Collections Report Recommends Path Forward

The Cloud Library project—jointly designed and executed by OCLC Research, the HathiTrust, New York University’s Elmer Holmes Bobst Library, and the Research Collections Access & Preservation (ReCAP) consortium, with support from The Andrew W. Mellon Foundation—has concluded after their year-long study of data from the HathiTrust, ReCAP, and WorldCat that:


There is sufficient material in the mass-digitized library collection managed by the HathiTrust to duplicate a sizeable (and growing) portion of virtually any academic library in the United States, and there is adequate duplication between the shared digital repository and large-scale print storage facilities to enable a great number of academic libraries to reconsider their local print management operations.

While the study confirms that academic libraries could have substantial cost savings by outsourcing the management of their redundant low-use materials, the authors acknowledged that “the organizational change required to achieve these gains is likely to be substantial and challenging to implement.” Particular obstacles identified were that: less than 20% of works are in the public domain, service providers will need to collaborate to obtain needed coverage, and better discovery and delivery are needed. The long-awaited Google Book Settlement is also an issue for determining how to license access to a shared e-collection.

If the HathiTrust Digital Library continues its current growth rate, the study projected that over 60% of ARL Libraries’ retrospective print collections will be in the repository by June 2014. HathiTrust doubled in size just during the 12 months of the study and is already larger than the average ARL library. Over 95% of the titles in the repository are books.

An analysis of one partner in the study, New York University’s Elmer Holmes Bobst Library, found that approximately 30% of the library’s titles were in HathiTrust, which represents 44,000 linear feet of library shelving that could potentially be removed. (Although even in the best of situations, a total removal of all of these titles is unlikely.) Library space is not the only savings identified. A high-density storage service like ReCAP is estimated to cost \$0.86 per volume to manage compared to \$4.26 in an on-site library collection.

The study report recommends several strategies that can be taken to advance the vision of a shared digital repository and large-scale print storage facilities and notes that: “This work will be challenging and deserves external support and endorsement by library leadership organizations and funders.” | NW | doi: 10.3789/isqv23n2.2011.10

 The complete report is available at: www.oclc.org/research/publications/library/2011/2011-01.pdf

If the HathiTrust Digital Library continues its current growth rate, the study projected that over 60% of ARL Libraries’ retrospective print collections will be in the repository by June 2014.





In Development or Revision

Listed below are the NISO working groups that are currently developing new or revised standards, recommended practices, or reports. Refer to the NISO website (www.niso.org/workrooms/) and the *Newsline* quarterly supplements, *Working Group Connection* (www.niso.org/publications/newsline/), for updates on the working group activities.

WORKING GROUP	STATUS
DAISY Revision Co-chairs: Markus Gylling, George Kerscher	NISO Z39.86-201x, Part A, Authoring and Interchange Framework Issued as DSFTU through September 28, 2011.
ERM Data Standards & Best Practices Review Co-chairs: Ivy Anderson, Tim Jewell	Technical Report in development.
Establishing Suggested Practices Regarding Single Sign-On (ESPreSSO) Co-chairs: Steve Carmody, Harry Kaplanian	NISO RP-11-201x, Establishing Suggested Practices Regarding Single Sign-On Recommended Practice issued for public comment.
Institutional Identifiers (I²) Co-chairs: Grace Agnew, Oliver Pesch	NISO Z39.94-201x, Institutional Identifiers Standard in development.
Improving OpenURLs Through Analytics (IOTA) Chair: Adam Chandler	Technical Report in development.
Knowledge Base and Related Tools (KBART) Phase II <i>Joint project with UKSG</i> Co-chairs: Andreas Biedenbach, Sarah Pearson	NISO RP-9-2010, KBART: Knowledge Bases and Related Tools Phase I Recommended Practice issued January 2010. Phase II Recommended Practice in development.
Physical Delivery of Library Resources Co-chairs: Valerie Horton, Diana Sachs-Silveira	NISO-RP-12-201x, Physical Delivery of Library Resources Recommended Practice to be issued for public comment in June.
Presentation and Identification of E-Journals (PIE-J) Co-chairs: Bob Boissy, Cindy Hepfer	Recommended Practice in development.
RFID for Library Applications Revision Co-chairs: Vinod Chachra, Paul Sevcik	NISO-RP-6-201x, RFID in U.S. Libraries Recommended practice issued for public comment.
Standardized Markup for Journal Articles Co-chairs: Jeff Beck, B. Tommie Usdin	Z39.96-201x, JATS: Journal Article Tag Suite Issued as DSFTU through September 30, 2011.
Supplemental Journal Article Materials <i>Joint project with NFAIS</i> Co-chairs Business Working Group: Linda Beebe, Marie McVeigh Co-chairs Technical Working Group: Dave Martinsen, Alexander (Sasha) Schwarzman	Recommended Practice in Development.



SUPPLEMENTAL JOURNAL MATERIALS

STRUCTURING CONTENT FOR USE

DAISY REVISION

Simplification, Broader Application Key

ANSI/NISO Z39.86, formerly called *Specifications for the Digital Talking Book* – more commonly known as DAISY, in recognition of the Maintenance Agency for this standard – has undergone revision to reduce complexity, improve and extend the user experience, support materials beyond the book (e.g., newspapers, audio tours, museum exhibits, presentations, and more), align with mainstream publishing, and allow for innovation. DAISY allows for content to be transformed into multiple output formats, including accessible formats such as Braille, DAISY DTBs, and large print. At this time, the *Authoring and Interchange Framework* is a draft standard for trial use.

JATS: JOURNAL ARTICLE TAG SUITE

Standardized Markup for Journal Articles

JATS (NISO Z39.96) provides a common format in which publishers and archives could exchange journal content. Based on the long-standing and well-accepted NLM Journal Archiving and Interchange Tag Suite, this standard defines elements and attributes that describe metadata and full content of scholarly journal articles. A draft standard for trial use is now available for testing. Three tag sets are included: Journal Archive & Interchange, Journal Publishing, and Article Authoring.

SUPPLEMENTAL JOURNAL MATERIALS

A Joint NISO/NFAIS Project

This project will recommend best practices for publisher inclusion, handling, display, and preservation of supplemental journal article materials. A Business Working Group is focusing on semantic and policy issues related to delivering materials that are supplemental to scholarly journal articles, while the Technical Working Group addresses issues such as metadata, persistent identifiers, linking mechanisms, packaging, and more.

Where to Get More information:

DAISY: Authoring and Interchange Framework

Co-chairs: Markus Gylling, DAISY Consortium;
George Kerscher, DAISY Consortium

ⓧ www.niso.org/workrooms/daisy

ⓧ www.daisy.org/zw/Main_Page

JATS: Journal Article Tag Suite

Co-chairs: Jeff Beck, NCBI, National Library
of Medicine; B. Tommie Usdin, Mulberry
Technologies, Inc.

ⓧ www.niso.org/workrooms/journalmarkup

**NISO/NFAIS Supplemental Journal
Article Materials**

Business Working Group Co-chairs: Linda Beebe,
American Psychological Association; Marie
McVeigh, Thomson Reuters

Technical Working Group Co-chairs: Dave
Martinsen, American Chemical Society; Sasha
Schwarzman, American Geophysical Union

ⓧ www.niso.org/workrooms/supplemental



NISO

www.niso.org



NISO
2-DAY
Forum

THE E-BOOKS ENVIRONMENT

OCTOBER 24 - 25, 2011

TREMONT PLAZA HOTEL | BALTIMORE, MD

Join us at the NISO Forum on *The E-Books Environment* as we explore distribution models, platform interoperability, and archiving and preservation issues from a variety of industry, scholarly, and consumer viewpoints. Participate in the community discussion to advance e-book development and support.

E-books have existed in the library landscape for over a decade, but it is only in the last few years that their use has grown to finally become the game-changer that many have anticipated for so long. E-book availability, distribution, licensing, discoverability, usage, and current and future access require content providers and libraries to adapt many of their existing processes. Amidst this chaos is a wealth of opportunities for new collaborations and initiatives.



NISO

FOR INFORMATION AND TO REGISTER, VISIT:

www.niso.org/news/events/2011/ebooks