

# E-books: Beyond the hype

## Essentials for scholarly and professional publishers

The topic of e-books has dominated the newsletters and social media networks covering the book publishing industry and, especially around the introduction Apple's iPad, the lay media seem to have expanded markedly their coverage of e-books and e-readers. Although e-book sales currently account for a modest few percent of book sales, the main reason for the attention is the expectation for rapid growth. Forrester Research estimates that 3 million e-book readers were sold in 2009 and 6 million will be sold in 2010. And these numbers do not include the tens of millions of laptops and smartphones that can display e-books.

One month after its introduction in April 2010, Apple announced that it had sold a million units, and that it had sold 1.5 million books through its iBookstore. Demand has been so strong in the U.S. that Apple has had to delay introduction of the device in other countries. Another factor behind the increased interest is the imminent entry of Google into the market. Google has announced that it will soon launch an e-book retail service that will enable customers to buy e-books from multiple sites and for multiple devices. Based on Google's history, a Google e-book site is likely to be user-friendly and therefore could expand the market rapidly.

For the scholarly and professional markets, publishers have taken notice of librarians' interest in purchasing ebooks as institutions continue their transition to digital collections. Jim Shanahan, editor-in-chief of McGraw-Hill Professional Publishing says simply, "My gut tells me that 2010 is a tipping point."

*My gut tells me that 2010 is a tipping point.*

*Jim Shanahan, editor-in-chief,  
McGraw-Hill Professional Publishing*

How to prepare e-book content has captured the attention of publishers given the proliferation of platforms. Amazon's Kindle catalyzed the e-reader industry and now faces robust competition from the B&N Nook, Sony Reader, MobiPocket (Amazon), Microsoft Reader, the two-screen eDGe from Entourage System, Apple's iPad, and multiple smartphone platforms such as iPhone from Apple and Droid from Google. (Fig 1) Each new platform so far has meant a different set of file specifications, assorted schemes for paying publishers, and various value propositions for the consumer.

e-book readers	Smartphone platforms	Multimedia tablets
<ul style="list-style-type: none"> <li>Kindle (Amazon)</li> <li>Nook (Barnes &amp; Noble)</li> <li>Sony Reader</li> <li>Microsoft Reader</li> <li>MobiPocket (Amazon)</li> <li>eDGe (Entourage)</li> </ul>	<ul style="list-style-type: none"> <li>iPhone (Apple)</li> <li>Palm</li> <li>Droid (Google)</li> <li>Blackberry</li> <li>Skype, Inc (medical)</li> </ul>	<ul style="list-style-type: none"> <li>iPad (Apple)</li> </ul>

Figure 1: Today's leading consumer e-book platforms

Multi-publisher platforms	Proprietary platforms	Book Wholesalers/Jobbers
<p><b>Comprehensive</b></p> <ul style="list-style-type: none"> <li>• Books@Ovid (Wolters Kluwer)</li> <li>• ebrary</li> <li>• NetLibrary</li> <li>• ProQuest</li> </ul> <p><b>Specialized</b></p> <ul style="list-style-type: none"> <li>• HighWire Press (scholarly)</li> <li>• Knovel (engineering oriented)</li> <li>• R2 Digital Library (health sciences)</li> <li>• Safari (technical)</li> <li>• StatRef (health sciences)</li> </ul>	<ul style="list-style-type: none"> <li>• Access Medicine/Surgery/Pharmacy (McGraw-Hill)</li> <li>• InterScience (Wiley)</li> <li>• MDConsult (Elsevier)</li> <li>• MedicinesComplete (Pharmaceutical Press)</li> <li>• Psychiatry Online (APPI)</li> <li>• ScienceDirect (Elsevier)</li> <li>• SpringerLink (Springer Science + Business Media)</li> </ul>	<ul style="list-style-type: none"> <li>• EBSCO</li> <li>• Ingram Digital (MyiLibrary)</li> <li>• Baker &amp; Taylor (YBP Library Services)</li> </ul>

**Figure 2: Examples of significant e-book platforms for institutions**

Beyond the individual consumer market, many publishers rely on institutional sales. Libraries are quickly embracing digital resources to provide information the way in which their clients demand—anytime, anywhere, anyhow—and as a way to save on premium space for their respective institutions. Publishers can provide digital versions on library platforms such as NetLibrary, ebrary, Ovid, MD Consult, or via jobbers such as YBP, Ingram, and EBSCO; and here, too, each vendor has different file requirements. (Fig 2)

by incorporating digital rights management (DRM) into their e-book initiatives. DRM is particularly challenging in the scholarly, technical, medical (STM) market as readers in this market often want to print pages for closer study or copy both text and images for use in other applications. If publishers deny these rights completely, users will be angry and sales may be adversely impacted; if publishers' policies are too liberal, publishers risk losing control of their content and may be unable to maximize their revenues.

## E-book acquisition budgets are expected to double in the next 5 years.

*HighWire Press librarian survey, 2010*

A recent survey of librarians from HighWire Press found that while e-book acquisition budgets are currently 10% or less of overall acquisition budgets in more than 90% of libraries, this proportion will roughly double in the next five years. Based on the consulting practices of The Lampert Consultancy and the Kaufman-Wills Group, growth in e-book acquisition budgets have come primarily at the expense of print book acquisitions, meaning, at minimum, a zero-sum game for publishers who are prepared but most certainly a losing proposition for publishers who do not embrace e-books.

Any discussion of digital book publishing must include the issue of copyright infringement and piracy. The publishing industry intends to avoid the mistakes of music companies

### Early days

The first e-books were the brainchild of University of Illinois student Michael Hart, who, in 1971, began posting famous texts like the works of Shakespeare in what became known as Project Gutenberg. Digital versions of books have been on the market since the mid-1990s. The first such products were CD-ROMs containing the full text of one or more books, with limited search functionality. Soon referred to as “shovelware,” they were designed for use by a single user although usually nothing presented the user from lending it to someone else. Moreover, despite the application of copy protection software, many users were able to make copies for friends and colleagues.

By the late 1990s, Web-based aggregates such as ProQuest, Ovid, and MDConsult appeared. These browser-based products were essentially shovelware on a larger scale, aimed primarily at libraries. Access and entitlement were administered by the libraries and the aggregates applied significant security within their servers, making this a much more secure environment for the intellectual property.

For the most part, these early e-book technologies provided only a generic HTML display. While convenient to view on screen, the format did not retain typefaces, page layouts, or other reader-friendly characteristics. HTML-oriented display is still used by a number of institutionally oriented informational platforms albeit with significantly improved search interfaces and other upgrades.

Over the past several years, PDF-oriented platforms have emerged primarily serving libraries and institutions. In most instances, PDFs alone are insufficient. Platforms such as ebrary and NetLibrary provide patrons with chapter metadata and libraries with metadata for MARC and OPAC systems. Chapter metadata can be supplied by publishers, often based on author input; author involvement, in principle, should enhance the relevance of search results and the discoverability of the content. Most publishers already provide MARC and OPAC records, so publishers can repurpose their current data. These supplemental data facilitate two critical aspects for libraries: 1) searching and 2) discoverability.

E-books for individuals advanced incrementally as broadband Internet access expanded, first with wired and then via wireless networks, and devices advanced from laptops to PDAs to smartphones. Amazon's Kindle, introduced in 2007, was the breakthrough e-reader. With its proprietary operating system, reflective eInk display technology, relatively simple "do-it-yourself" file uploads for publishers, and user-friendly e-commerce on the Amazon platform, the Kindle quickly became the device to beat. Other e-book readers—B&N Nook, Sony Reader, and the like—provide reading functionality similar to Kindle's along with modest enhancements. An unusual device, the eDGE, offers two screens, an eInk screen similar to Kindle's as well as an LCD similar to a typical laptop. The LCD displays images and multimedia in ways that the eInk screen cannot.

Apple's iPad could represent a fundamentally different approach. Although it excels in some respects as an e-book reader (including a Kindle app), the iPad runs an operating system with broad multimedia capabilities and an ever-increasing number of applications that go well beyond e-books. Conceivably, then, the iPad platform could be configured to provide seamless connections between e-books (Kindle, PDF, etc.) and multimedia, functionality that has caught the attention of higher-education publishers among others.

### Production basics

Newsletters and blogs are filled with speculation about which e-reader platforms will emerge as leaders over the next couple of years, not to mention which additional platforms will enter the market. Publishers are not in a position to choose sides; they need to consider e-book availability for whatever platform(s) their customers prefer. For instance, American Psychiatric Publishing, Inc. (APPI), the publishing arm of American Psychiatric Association, offers e-book titles on multiple platforms, according to Bob Pursell, associate publisher for advertising, sales and marketing. These include APPI's proprietary platform Psychiatry Online; library platforms ebrary and NetLibrary, and Ingram's MyLibrary; the retail oriented ebook.com; for e-readers Nook (at present), eDGE (coming soon), and other devices.

For publishers, providing content to multiple platforms raises two key production challenges—1) versioning and 2) conversion.

Solving the versioning problem involves a fundamental shift in the publishing paradigm. Just as the online version of journals has become the version of record, no longer is the print book the primary product, with e-books as products derived from the print book. Instead, print books and e-books need to be two (or more) varieties of output from a single, definitive data file. (Fig 3)

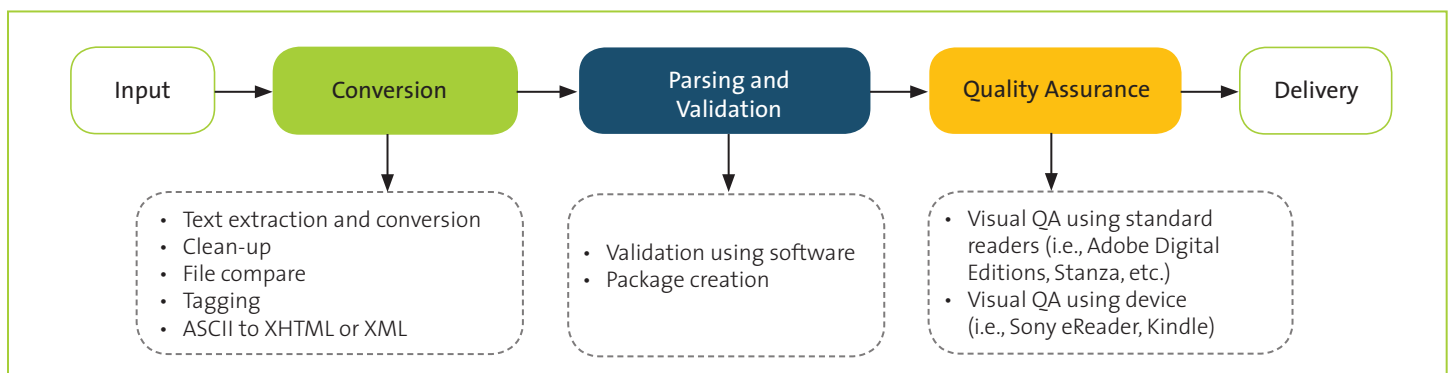


Figure 3: Example of file conversion workflow

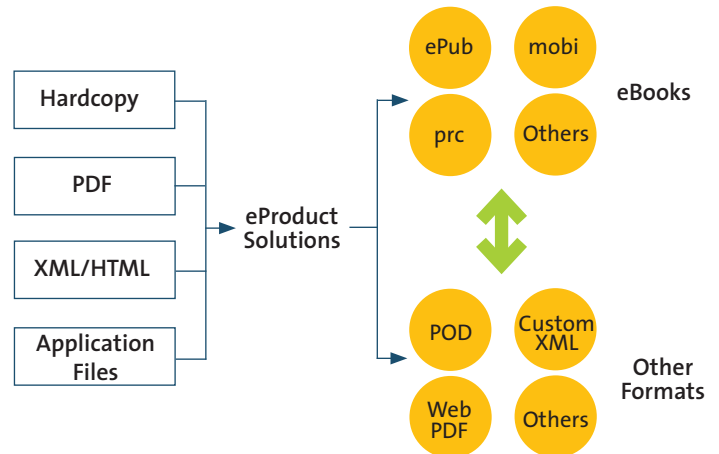
For most publishers, the most likely solution is XML, preferably using a DTD across a wide-ranging product line rather than a single product, and preferably introducing XML at the earliest possible production stage.

Publishers need to adjust their workflows so that all revisions and proofreading are made to the master XML file rather than to a book PDF or some other derivative. The other challenge, which occurs frequently when publishers repurpose backlist titles, is simply identifying and archiving the single most authoritative version of the content file. Looking forward, publishers should either invest in a robust Content Management System (CMS), or at least develop a protocol to identify and archive “the” XML file for each book title. With an authoritative file in hand, converting well-structured XML files to most e-book platforms, while not quite automatic, is usually an orderly and relatively inexpensive process.

Jabin White, Director, Strategic Content at Wolters Kluwer Health, says that the biggest challenges in implementing this forward-looking approach are cultural. “We are asking publishing people to recognize the value of this systematic approach. My goal is to make this so publisher-friendly that it’s easy and natural for publishers to use even under deadline pressure.”

Converting backlist titles to e-books presents a different set of challenges. In the words of Jabin White, “There’s a different story for every title.” At many publishers, composition files for backlist titles often exist in a variety of formats. For instance, text files may be SGML – or, more problematically, SGML with on-the-fly formatting commands to adjust the appearance of type without writing additional specifications. Tables are a particularly knotty challenge in terms of usable coding. Graphics files may or may not be stored with text files, different figures may be saved at different resolutions, and file formats may be different. Within STM disciplines, some publishers may be grappling with TeX or LaTeX, which are very efficient for setting mathematical notation but can be challenging to convert to XML. Some small publishers may generate their content within desktop publishing software such as Quark Xpress or Adobe InDesign. Each presents its own conversion challenges depending on the files themselves as well as the destination platform.

Although file conversion is sometimes managed by publishers’ production departments, all aspects of the workflow are different. Given the tight staffing in most publishing organizations, it usually makes sense to turn to an outside vendor that can work with multiple file types and produce content fit for multiple platforms and for multiple uses. (Fig 4)



**Figure 4: End-to-end digital product solutions**

Even once outsourcing has been adopted, publishers need to have an in-house point person to track down and verify files and ensure that the contractor has all the necessary output specifications. When thinking about the investment implications, publishers should realize that converting backlist files to XML will probably facilitate additional repurposing, potentially broadening the future revenue stream beyond that of current platforms.

Ironically, Kindle, the platform that has defined a large segment of the e-book market, can be a significant challenge for many STM, professional, and higher-education publishers because of the device’s built-in display limitations. Kindle displays text very effectively, albeit without reference to a publisher’s own typography. Amazon’s Kindle Store carries thousands of STM titles, so clearly publishers and users do cope with the device’s limitations. However, both figures and tables are difficult to display on Kindle, and more complex ones may be unusable. And book titles that rely on two-column designs or marginal callouts won’t display those layouts effectively, particularly on the relatively small original Kindle.

Publishers, as well as device makers and software houses, are eager to automate conversions from hardcopy books to e-books as far as possible. The most viable standard at this time is EPUB, an open standard that builds on familiar systems such as XML and CSS (Cascading Style Sheet). The EPUB standard is an effort of the International Digital Publishing Forum ([www.idpf.org](http://www.idpf.org)) that has drawn support from organizations such as Association of American Publishers. While EPUB is relatively robust when dealing with text-heavy content, it can stumble when confronted with some of the familiar challenges – complex page layouts, heavy graphics, and the like. Even when EPUB is effective, new users face a significant learning curve. To date, EPUB represents a constructive effort, and EPUB devices – notably the e-reader apps on the iPad – are beginning to reach consumers.

At this time, publishers need to rely on vendors’ ability to recognize challenges related to file formats, content complexities, and limitations of e-reader devices themselves. It takes a combination of experience, staff resources, and careful quality control to ensure that the user experience works. Also, a large, diverse vendor operation can overcome technical challenges while still getting products to market in a timely way.

### Digital Rights Management

As a practical matter, DRM boils down to a short list of simple questions. (Fig 5)

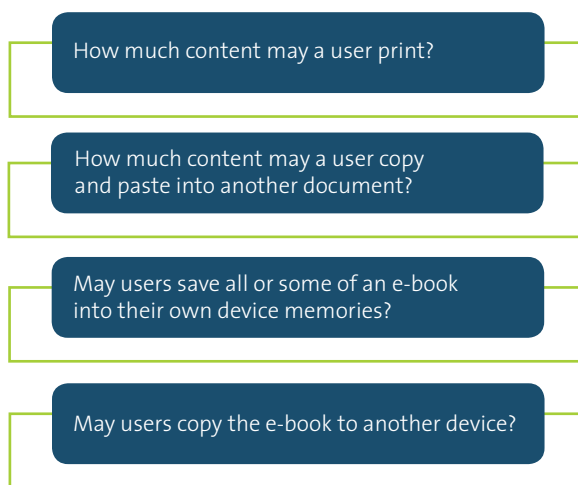


Figure 5: Set of questions to help determine appropriate level of DRM

There is no one set of right answers to these questions. Publishers need to answer the questions based on their understanding of their own markets. For instance, many STM researchers are more comfortable or prefer paper copies of materials. In the survey of librarians by HighWire, the vast majority of respondents identified limitations on printing as a “significant” or “very significant” issue.

To a degree, DRM considerations may determine which e-book platform/s a publisher supports. For instance, a publisher that is concerned about clients holding any e-books on their own hard drives may choose a platform that simply serves e-book content out of a dedicated server that the publisher or aggregator controls. On these platforms, users establish an account and then obtain access to specific titles based on whatever e-commerce processes the publisher has established. In contrast, other publishers may choose to download files into the user’s device and establish DRM within the downloaded file.

Some information providers have decided to bypass DRM and trust their customers to do the right thing. One such organization is National Science Teachers Association. Tyson Brown, their director of new products and services, says, “We have a nagging fear that by not utilizing DRM, we are exposing ourselves. However, teachers generally want to do the right thing. Sometimes they even ask our permission for uses that are clearly within the bounds of Fair Use under copyright law. We want our content to be as widely available as possible.” And NSTA is monitoring sales patterns, hoping to detect and combat piracy if it develops.

No matter what DRM a publisher applies, it is prudent to follow NSTA’s practice of monitoring the Web and social networking sites on a regular basis. One simple approach is to set up a Google Alert—essentially a Google search that runs automatically at a specified frequency—for identifying characteristics such as author name, book title (if reasonably distinctive), or maybe even ISBN. A third party that is trying to make money from selling a pirated e-book probably will make the material available through a Web site, and the Google Alert should direct a publisher to the offending URL. Publishers also should search periodically in Facebook and Twitter.

Author	Title	Kindle price	Print list price	Kindle print discount
Rhodes	Crystallography Made Crystal Clear	\$38.45	\$60.95	37%
Blow	Outline of Crystallography for Biologists	\$64.83	\$85.00	24%
Borchart-Ott	Crystallography	\$84.00	\$105.00	20%
Prince	Mathematical Techniques in Crystallography and Materials Science	\$55.96	\$69.95	20%
Ferraris	Crystallography of Modular Materials	\$56.00	\$70.00	20%

**Table 1: Result of simple search in Amazon's Kindle store**

### Business models

Most publishers and book buyers know that Amazon purchases Kindle “copies” at 50% of a publisher’s list price but sells trade book best-sellers at \$9.99 regardless of the publisher’s list price. For a \$27.95 hardback trade title, then, Amazon pays \$13.97 and loses \$3.98 on every Kindle sale. The introduction of the iPad is shaking that model. Apple has reached so-called agency agreements with most of the dominant trade publishing conglomerates, which permit publishers to establish a selling price. In return, Apple takes 30% of the revenue. Publishers seem to be moving toward a \$14.99 list price for trade best-sellers for the iPad.

This high-profile controversy has obscured a much different truth in markets other than trade bestsellers. Table 1 shows the result of a simple search in Amazon’s Kindle store. “Crystallography” was chosen for our search term, and the table shows information for the top five titles returned by the search.

The lesson here is clear and applies across thousands of STM titles: Kindle prices are based on print prices with relatively modest discounts. If file preparation costs are low, Kindle sales return contribution margins that can actually exceed margins on sales of print copies. In fact, discounts are comparable to the ones publishers give to brick-and-mortar stores on similar titles. Furthermore, Kindle involves no PP&B costs and no inventory risk.

Business models for libraries are more complex. Models now in use include perpetual access (the closest approximation to buying and owning a print copy), concurrent users, annual license, purchase with annual maintenance, and others.

As with periodicals, librarians also expect significant services from the platforms supporting the e-books including meaningful usage data, easy integration into MARC and OPAC, and detailed metadata to enhance discoverability.

As argued earlier, publishers need to rethink their production processes to focus on the information per se rather than the print book format. When considering pricing, publishers need to make an analogous shift. Instead of thinking about a price tag on a print product they hold in their hands, publishers need to think about the value of the information to an individual reader and to a group of library patrons. In some cases, as for the crystallography titles in Table 1, publishers (and presumably individual customers) have decided that the value proposition for a Kindle version is very similar to the print version.

---

Based on the HighWire survey, librarians greatly prefer the perpetual access model with prices closely related to print prices but are willing to consider other models.

---

Many STM and professional publishers set prices for individual chapters of their e-books based not on the list price of the complete e-book but on the value proposition for the chapter.

The thought process about consumer pricing will become more complex as the iPad and other multimedia-oriented tablet products gain market share. Consumers are likely to expect that an e-book on one of these platforms also will take advantage of the multi-media capabilities. As Jim Shanahan of McGraw-Hill pointed out, his customers will expect a medical e-book to include videos and animations that don’t exist inside the print books. Should the iPad version be priced higher than the book? If so, by how much? Shanahan admits that his company, like many others, is in the very early stages of thinking about this.

---

## Many STM and professional publishers set prices for individual chapters of their e-books based not on the list price of the complete e-book but on the value proposition for the chapter.

---

Most book publishers practice multi-channel marketing, aiming to reach individuals, the book trade, and libraries throughout the world with every customer having multiple ways to discover and buy their products. Publishers are applying similar principles for distributing e-versions of their books. Selecting e-book publishing alliances is more art than science.

Bob Pursell mentions two key criteria: 1) “attractive terms” and 2) “the highest level of discoverability.” Often the strategy involves both proprietary and third-party platforms. For instance, APPI offers about 250 of its titles via multi-publisher content aggregators while some of the publisher’s top-selling titles remain exclusive to its proprietary platform. Pursell says, “For now, we wish to provide those titles in a semantic search environment exclusively at Psychiatry Online. To gain a foothold in the market and gain larger revenues, we feel our strategic direction is to deliver e-books ourselves.”

On the other hand, Tyson Brown of NSTA points to the technical advantages of allying with third-party platforms. Platforms provide institutions with guarantees of 24/7 availability, they administer access and entitlement, and they interface safely with institutional IT systems. The other key advantage of multi-publisher platforms is the power of searching within a large and diverse collection, which in turn can increase sales and usage of a publisher’s titles.

### Setting priorities

Readers involved in the book publishing industry know firsthand that there’s too much to do and not enough people in-house to do it. To help determine priorities, book publishers may wish to consult this list of critical issues regarding e-books:

### Commit to strategic investments

Production initiatives such as file conversion and the development of XML DTDs take money. Hard-pressed book publishing operations may be reluctant to erode their bottom lines for a market that is still a tiny fraction of the market for print. However, Lisa McAllister, Director, Product Strategy at Wolters Kluwer Health, points out, “There’s a long time lag in medical publishing. Publishing decisions we make in 2010 relate to books we’ll publish in 2012.” As a result, she suggests that publishers retool their book proposals to account for both anticipated revenues and anticipated costs related to e-book demand a couple of years into the future.

### Assess your backlist.

Which titles have ongoing sales potential? For those titles, where are the composition files, and are they configured for efficient conversion to e-books? Seek expert advice on the technical challenges.

### Organize production of frontlist and future titles to facilitate digital formats.

The goal is fully XML-coded content contained in a robust Content Management System. In the real world, a step-by-step approach will probably work best and partnering with outside experts and contractors can amplify the skill set of in-house personnel.

### Talk to both institutional and individual customers.

Focus on e-book publishing on platforms that matter most to them and not necessarily on every available platform.

### Manage expectations.

Remember that e-books currently account for a small share of the book market. Adding e-books to a publisher’s product mix is unlikely to generate a big spike in sales. However, failing to add e-books looks like a risky proposition for the future.

## Contributors

Special thanks to the following individual contributors:

- Tyson Brown, Director, New Products and Services, National Science Teachers Association
- Lisa McAllister, Director, Product Strategy, Wolters Kluwer Health
- Bob Pursell, Associate Publisher, Advertising, Sales, and Marketing, American Psychiatric Publishing, Inc.
- Jim Shanahan, Editor-in-Chief, McGraw-Hill Professional Publishing
- Jabin White, Director, Strategic Content, Wolters Kluwer Health

## The Authors

- The Lampert Consultancy - **Rich Lampert**  
[www.lampert-consultancy.net](http://www.lampert-consultancy.net)
- Kaufman-Wills Group - **Cara Kaufman**  
[www.kaufmanwills.com](http://www.kaufmanwills.com)

**Rich Lampert** is owner of The Lampert Consultancy, LLC, established in 2004 to provide strategic, editorial, and marketing services to publishers in STM, professional, and scholarly publishing. Rich is also, Principal, Publishing Services Division, at Doody Enterprises, Inc., which focuses on not-for-profit publishers.

**Cara Kaufman** is co-founder of Kaufman-Wills Group, LLC (KWG), which was created in 2000, to offer STM and other scholarly publishers a full range of professional publishing services in the areas of strategic planning, business development, electronic publishing strategy, RFP and self-publishing projects, editorial services, and marketing and market research.

SPi sought the help of KWG in developing this white paper.