Editor’s Note: This article was originally published in eContent Quarterly (March 2014). The e-book market is volatile. Many components of the pricing analysis have changed, some for the better, some not.

Project Gutenberg began creating e-books in 1971, but e-books did not gain any real traction until 2007, when Amazon launched the Kindle. Even in the late 1990s, when e-books were at least on the edge of librarians’, academics’, and publishers’ awareness, it seemed that e-books never would be embraced by the general public. Ultimately, they were, and the questions facing us now are not how viable e-books are, but how much of the book market will be dominated by e-books and whether e-books will replace print books.

This should not have been a surprise, as the music market saw widespread disruption in the 1990s when early e-readers like the Rocket eBook were floundering. Clearly, consumers were interested in the obvious benefits of e-content, and while Sony launched its e-reader first, it was Amazon, the inventor (practically speaking) and leader of online retail, who realized that ease of use and delivery of content would make or break the e-book market. Again, this simple design solution should have been obvious as early peer-to-peer services like Napster and WinMX had proved consumer interest in e-content if a simple enough interface was available. Librarians and information scientists should not have been surprised, as the behavior patterns and adoption of e-content fit nicely within well-established information science and information-seeking behavior. Instead it was Amazon and Apple that met the needs of patrons with a viable electronic ecosystem for their music, video, and even books.

Public librarians were surprised and totally unprepared for the explosive growth of the e-book market in spite of the clear evidence from the music and movie industries. This growth and adoption of e-content was predictable, as information professionals had been subscribing to digital services like EBSCO and Gale for years. Unfortunately, the profession seems to have assumed that those types of e-content would never replace physical content, and would remain a supplemental service to our core service as physical content warehouses. Evidence suggests that librarians finally are starting to realize the bigger-picture issues involving the first-sale doctrine and its lack of application to digital formats, as well as the possibility that digital formats may replace physical ones. The involvement of librarians in new organizations like the Owners’ Rights Initiative and several high-profile lawsuits involving (mainly) academic libraries have begun to capture the attention of librarians everywhere. While Art Brodsky’s celebrated (and problematic) article in Wired does not mention first sale specifically, it discusses the core problems surrounding a lack of first sale for e-content.

The Historical Context

The pressing question facing the profession is how much physical content will be exclusively available electronically, and when will that happen? For public librarians, especially, a third question is how it will affect their budgets. This article will answer those questions using an experiment by Mesa County Public Libraries, Colorado, to forecast a number of conservative assumptions about an eventual migration to e-content, and apply it to the best assumptions we can make about current e-content costs. We chose to use e-books primarily because print books represent the core of our purchases and physical spaces. While acknowledging that databases can and most probably will continue to absorb more of our budget, we
focused on books exclusively, because a transition to digital books would represent the most chaotic transition for us as institutions given that so much of our physical space, human resources, and public perceptions still revolve around print materials. However, in order to understand the forecast, a broad examination of the e-content market and ecosystem is needed.

It is imperative to recognize that the e-content market is in its early stages of development. Any examination of it must be understood in this context. For example, the much-celebrated Pew study published in early 2012—which found library e-book borrowers also buy e-books—lacked any real examination of the motivation for purchasing e-books. If the observation about ease of use and interface design is correct, it may be that e-book borrowers are bypassing the demonstrably clunky and click-intensive interface of systems like OverDrive and EBSCO and buying the e-book version out of frustration rather than putting up with subpar, buggy interfaces. ALA President Molly Raphael's comment that “e-book borrowers being buyers is a phenomenon that's true in the print world as well” is a non sequitur. It is a correlation that does not have demonstrated causal links. Rather, she and the rest of the profession are carrying an assumption over from the old physical market into the digital one. Given how spectacularly the profession failed to predict and understand the e-content impetus and its explosive potential, her assertion is suspect at best. The Pew study is best understood and utilized as one snapshot in the chaotic and explosive evolution of a market, rather than a foundational and authoritative examination of said market. Doing otherwise is akin to using an australopithecine as a stand-in for Homo sapiens. In other words, don’t assume the score at half-time will reflect the final score.

Music is the most mature e-content market and the best one for extrapolating and forecasting potential trends for the rest of the market. This is because we can confidently assume that the last physical format for music, CDs, is in its twilight stage—soon to be replaced entirely by digital formats and niche markets like LPs. While this is common knowledge, it is less remarked upon that digital sales have yet to fill the void left by the decline in CD sales and revenue. Music’s “lost decade” is a reference to the period of time between 1996 and 2006 when revenues for the music industry went over a cliff. In February 2011, Michael Degusta argued in Business Insider that music revenues were down 64 percent from their peak in 1999. While various estimates differ on the actual amount of lost revenue, the implications were clear: the music industry had been savaged by digitization. Like all statistics, these numbers may be somewhat misleading. For example, the peak period of time referred to may have actually been a bubble in revenue due to the low production costs of CDs, which were sold at a high profit. Regardless, for the industry and the people who relied on it for their livelihood, the impact was real, significant, and is still felt today.

While the sale of singles had a growth curve, it was nowhere near enough to overcome an aggregate loss. Starting in the late 1990s, the music industry experienced a downward sales curve that continued until 2012 when music finally posted an increase in revenue thanks to a growth curve in digital sales (http://www.ifpi.org/content/library/dmr2012.pdf). The growth came in at an anemic 0.3 percent, but it was the first increase for the industry in thirteen years. For reasons already mentioned, this growth must be interpreted cautiously, but a few broader conclusions can be extrapolated from it. From the perspective of librarians who have made the shift from e-content skeptics to nervous anxiety about what trigger event will push content into exclusively digital formats, it should be clear that publishers and Hollywood are afraid of the profit loss the music industry experienced. This seems like a reasonable inference given the catastrophic loss of revenue their music-industry counterparts suffered.

For publishers, it is a careful balancing act of weighing various risks against expected benefits. For example, publishers can reasonably expect to see an increase in total revenue from digitally exclusive print runs even if total sales decline, as almost all of the original investment risks do not apply to e-content. Gone are the analyses, reports, and salaries dedicated to developing a reasonable “print run.” Gone are the production and distribution costs. Instead, publishers are looking at marketing and visibility as brick and mortar bookstores continue to lose visibility and market share. As Mike Shatzkin has correctly noted (http://www.idealogue.com/blog/marketing-will-replace-editorial-driving-force-behind-publishing-houses), marketing e-books is the critical question publishers are attempting to solve in their attempts to avoid the down curve in revenues the music industry experienced. Regardless, it seems clear that publishers and other content providers/producers are attempting to learn from the music industry and solve some of these questions prior to committing to e-content markets. They know from experience that once digital content is released into the marketplace they lose a certain amount of control of that content regardless of the protective measures taken. Digital content is inherently uncontrollable because copyright law evolved in markets where the reproduction side of the equation involved work and cost. Not so with digital material.

Currently, the first-sale doctrine does not apply to e-content either, which is a much-envied position for copyright holders. Librarians have naively believed for years that publishers were more or less willing partners rather than recognizing the first-sale doctrine as the foundation of our practice and services. Publishers
always have been skeptical of the evidence that print borrowers are also buyers. Without the first-sale doctrine, they are relatively free to shift the balance of power in their direction. Most consumers and librarians are blissfully unaware that they do not “own” any of the digital materials they purchase. The issue is further complicated as vendors like Amazon tend to use the same language on their websites as they do for physical materials. The infamous little orange button says Buy rather than License. Attractive buttons and icons require less time and energy than cumbersome End User License Agreements (EULAs). Yet, every user of legal e-content has agreed to them, most without knowing what they are. EULAs are those cumbersome, small-font, wordy boxes that pop up every time a user installs or sets up a new device/app/vendor. It is the Agree button for text that no one knows, reads, or understands in spite of the clear request to read in the first line below.

Most EULAs contain two main components:

1. A liability clause
2. A license clause

There is a lot of other legalese, but for consumers those are the two most critical components and the most attractive to copyright holders. Because even though all the other language and experiences suggest or imply that the e-content “purchased” is owned by the user, in reality the EULA waives first sale and agrees to a much more restrictive license. Currently, this is how the vast majority of e-content is regulated and distributed.

It is critical to note that none of this developed in a vacuum. Copyright law has long tails, and efforts to reconcile it with the digital world have been ongoing. It became a public issue in the late 1990s, and the Digital Millennium Copyright Act (DMCA) was the first major attempt to address the issue. What is notable is that the DMCA was written in the social, economic, and legal context of peer-to-peer file sharing. File sharing was at the forefront of everyone’s thinking, and the DMCA and subsequent thinking focused on protecting the copyright holder’s rights rather than other questions that could have taken precedence. It is possible that without piracy as the driving force, a broader perspective might have been taken. This seems reasonable given the 2001 Executive Summary by the United States Copyright Office on the concept of “Digital First Sale.” The summary explicitly acknowledges that technology has the capacity to potentially rectify one of the problems with e-content. That is, it can be “reproduced flawlessly” with little effort, placing it far outside the original boundaries envisioned by first sale. The legal doctrine is complicated, but it is easy to understand the situation previously described where publishers do not consider print runs with digital versions, because they are replicated on demand. First sale was developed for a far different market where production costs create real scarcity and physical items can reside with only a finite number of owners. The only way to expand the number of owners would be to replicate (at significant expense) the original item, which is subject to all the rules and regulations of the market.

The 2001 report, while dense, clearly communicates that the concern is not with technology’s ability to deal with this replication problem, as even the report acknowledges the possibility of a “forward and delete” technology that could ensure the original file leaves the owner’s possession and is transferred to another owner. Rather, the overall concern is the ability of the digital economy to ensure scarcity and control distribution—specifically illegal distribution. ReDigi, the intrepid start-up that attempted to create just such a forward and delete mechanism, recently discovered nothing has changed since 2001 as it lost the initial suit. It is this scenario and these fears that drive the library e-content market, and that will continue to drive it.

Can libraries reasonably forecast and plan for the future given the current uncertainties? Yes and no. Yes, because they can extrapolate from the music industry and expect content producers like Hollywood and publishers to fight for a legal environment that favors their interests. The epic eighteenth-century battles in England over copyright law are quite instructive in that we see the same two arguments over copyright restrictions being made today by the same parties. Copyright holders argue that loosening copyright will result in less production because it will remove motivations for artists to produce, while the other side argues against the dangers of monopoly and advances a broad ethical concept of “the public good.” So libraries can draw a clear line from OverDrive’s famous 700 percent increase of 2010 to the state of Kansas’s fees to Wiley’s recent announcement to limit downloads of articles to “100 full-text article/chapter/encyclopedia entries per day based on the previous day’s usage.” HarperCollins’s twenty-six-loan cap for e-books is best understood in the same light. Libraries should not be surprised anymore, as this type of behavior is to be expected from copyright holders fighting for their own existence and profits. If we add to this the lack of first-sale protection, we can begin to consider what the future may look like if the players with power, money, and influence win the battle.

Forecasting E-content in a District Public Library

At Mesa County Public Libraries (MCPL), we decided that there was enough data to project several different forecasts given a number of assumptions we could
include. Based on the reasoning presented above, it seems sensible to assume that all content producers are motivated to migrate content to digital formats exclusively. Reference materials and pulp fiction are the most obvious choices, as the value of these materials is related to their content and little value is placed in the format or medium. We have seen this most famously with the end of Encyclopaedia Britannica’s print editions. Britannica is simply emblematic of the larger trend, and is useful only in that it forced many doubters to acknowledge the potential reality facing libraries. It is safe to assume that content producers in all formats are looking at migrating additional content to digital formats exclusively, as they can increase profits and retain greater control of copyright.

Given the examination above, it may seem reasonable to start with our music collection, but we opted against that genre for numerous reasons. First, music does not represent a significant part of our collection and is less critical to patrons than books and movies. Second, library music collections are being rendered irrelevant by online services like Spotify and Pandora. Indeed, some urban libraries are already reporting drops in CD circulation that could be the first signs of the death of library popular music collections. But even in libraries like MCPL where CDs still move, they are not a significant part of our circulation, which corresponds well with libraries across the nation. As has already been argued, the music industry is the most mature digital market and serves as an excellent indicator. Because libraries have been so slow to respond to digitization, it is probably too late for libraries to develop plans for digital music given how crowded the market already is and how low the cost is for consumers to enter that market. There is much more that can be said about this issue, but it is outside the scope of this article. Suffice it to say, music is not a good genre for our forecast.

While DVDs represent a significant part of our circulation, we decided to use OverDrive and limit the forecast to e-books. Books represent roughly 50 percent of our circulation in comparison to DVDs, but there are bigger and more compelling reasons for using books as the key indicator. Books loom larger in the image and identity of libraries, and our past, present, and possibly future are deeply tied to their existence. The vast majority of our physical space is dedicated to books, and a disruption in this market threatens our services and identity in ways a loss of DVDs simply cannot match. In spite of all the activity and energy focused on developing new iterations of libraries as place, libraries as services, libraries as outreach, our identities and services remained anchored in books. Specifically, in the physical format of books. This is why e-books continue to dominate our discussions and command our attention, in spite of the clear need for a broader focus on e-content. We were also conscious of this trait among our colleagues and realized that forecasting trends for e-books would have a much greater impact on our colleagues than any other format.

Our first problem was that the market is so volatile. Costs have not been as stable as we would like, which is understandable given the nature of the evolving market. We also have emergent models like the Douglas County “ownership” model that are turning the standard model upside down. The question was how to develop a price model moving forward. Given this instability, the safest route was to look at movie licensing for streaming, which is another model that has enough maturity and stability to provide several clues. First, the industry seems to assume some kind of annual cost model. At its simplest, this takes the form of paying x dollars per year per title during the length of the contract. Even in contracts where an entire catalog is licensed this can be the underlying cost model. It is not always, but it seems that the industry finds this a satisfactory model. It is critical to note the difference between organizational licensing and individual licensing. While Netflix or Amazon Prime costs a few dollars per month per user, Amazon and Netflix are gambling that they can distribute their massive annual licensing fee as an organization across millions of users. For most libraries this is an unsustainable model for video content. Only libraries with millions of patrons in their tax base could even entertain such a model. Libraries are too limited geographically and numerically to distribute these types of licensing fees across millions of users like Netflix does. Moreover, as its shrinking digital catalog suggests, this may not be a sustainable model for Netflix either. It remains to be seen if Netflix can continue to offer enough content at its current consumer price to have long-term viability. Also, one of the most commented-upon effects of digitization on distribution markets has been the elimination of traditional middlemen, and Warner Brothers’ decision to remove its entire catalog from Netflix evinces a growing awareness on the part of copyright holders that marketing is more important than distribution.

Assuming that e-book licensing will eventually settle into an annual cost per title for libraries, the next relevant question is what those costs will look like. Given the behavior of publishers, notwithstanding Macmillan’s and others’ willingness to develop more library-friendly models, the behavior of the Big Five publishers indicates an interest in raising the cost per year above the cover price of the print version. While such a suggestion is anathema to librarians, there are legitimate reasons for this model. First, as has been noted previously, there are reasons to doubt the Pew survey that suggests e-book borrowers are also e-book buyers. The more mature digital markets indicate a commodification trend for items that were previously not treated as commodities. The simple act
The images in figure 2.1 are taken from a pricing comparison Douglas County has conducted the past few years for bestsellers. The goal is to draw awareness to the discrepancy in pricing between e-books and print books. The highlighted field is quite shocking as e-books are generally five to six times as expensive as their print counterparts. The problem is that the numbers are difficult to pin down as various libraries have individual contracts with the vendors and pay different hosting and maintenance fees. MCPL recognized the need to create an annual cost for an e-book, because budgets run annually and that number is the one needed to forecast the impact a shift to digital would have on our budgets and collections. We factored in a wide array of variables before developing the $35.85 number reported above. We looked at the average shelf life of their print counterparts, figured in the twenty-six-loan limit for HarperCollins, averaged out or related the cost of lower-circulating items to those that have high circulation, and other variables. Another challenge is that the factored cost per title for hosting and maintenance will actually go down as more titles are purchased. Neither were we able to factor in any kind of reduction in cost for technological advancements, as we have seen little movement away from Adobe Content Server (ACS), and the annual fees for ACS have not been stable the past five years. While we were not able to develop a precise formula for calculating the cost, we eventually settled on $35.85 as a reasonable cost per title per year given current trends and costs. Of course, given the volatility of the market described in the first part of this article, this number could change at any time. What is critical for librarians to understand is the need to begin assessing the actual cost of e-content on an ongoing annual basis, as similar pitfalls are embedded in services like Freegal (music) and Hoopla (pay-per-circ video streaming).

While the initial estimated cover price of e-books is a shock, the real cost is incurred at the annual level.

of digitization is having a significant effect on consumer behaviors and attitudes about that content. What the ultimate result will be remains to be seen, but we have ample evidence suggesting that an ecosystem awash in content is an ecosystem that drives content value down. If consumers begin seeing books as commodities, their motivation to purchase is significantly reduced, provided libraries offer easy access to digital versions. In such an environment, the fears of publishers regarding a negative effect on overall sales is at the very least a reasonable one, justifying a higher annual cost per e-book rather than a lower one.

Over a typical twelve-month period, MCPL pays an average price of $17.98 for any type of print book. The American Library Association does excellent work tracking public and academic library expenditures. Without delving deeply into the data, it is critical to note that MCPL is neither at the high end of the expense curve nor at the bottom, but is fairly representative of the “average” public library. Academic libraries probably report different numbers due to smaller print runs for scholarly materials, but it is safe to assume that the numbers have not shifted drastically at any point in the past few decades. It also is safe to assume that the average price of a print book has steadily crept upward, and that all libraries have been able to make their collection-development budgets work and meet the needs of their patrons within reason. This is not to minimize the pressures libraries have reported in recent years as budgets shrink and costs rise. Rather, it is to set up the stark contrast between challenges in the print age and challenges in the digital age. Using OverDrive’s current costs and assuming some of the above forecasting, we placed the average cost of an e-book at $35.85 annually for Mesa County. This is quite different from neighboring Douglas County Libraries numbers in their ongoing reports on average costs of e-books vs. print (see figure 2.1).
Library budgets have developed in a book market where the first-sale doctrine protected libraries from annual license costs. By using our collection development statistics and annual budgets from the past, we were able to calculate a very accurate average cost per book over a twelve-month period at $17.98. We are extremely confident in this number, because like most libraries we keep meticulous records and calculating that number was fairly straightforward. It means that over a ten-year period we purchased about ten different books for $179.80. But in the new annual licensing model, assuming that our estimated cost of $35.85 per year per title is correct, the $179.80 above would pay for only one title for five years. It is easier to visualize than read:

**Under first sale:**
$17.98 \times 10 = \$179.80$ (for ten different books over any given ten-year period)

**Under annual license model:**
$\$179.80 \div \$35.85 = 5$ (five years for one title)

Take the $\$179.80$ spent on ten books over a decade, and divide it by the estimated cost per year for one title, and the real problem surfaces. Libraries would be spending the same amount they normally spend on ten books in ten years on a single book in five years.

If the digitization “trigger” event described earlier occurs, or if the same natural migration happens the way it happened in the music (and potentially video) industry, print books slowly will begin to cease production and be replaced by digital-only versions. This is the question on almost everyone’s minds, and it is far beyond the scope of this article. Indeed, we could fill this entire journal issue with articles devoted to the subject and not have done it justice. For reasons already stated, discussion of the “death” of print may be premature, but at the same time it remains a very real possibility. Assuming that this possibility were to take place at a 5 percent migration rate, MCPL would start to see a significant decrease in materials inventory almost immediately. By “5 percent digitization rate” we mean that we are assuming that publishers migrate 5 percent of print books to digital formats exclusively each year. We used a 5 percent digitization rate because it is conservative and resulted in catastrophic results. We also capped it at 20 percent simply to make our calculations easier, and stopped at year nine because the final results were shocking enough without moving to year ten. Because budgets have remained relatively flat over the past five years, we assumed a relatively flat budget over the forecast period. This is a critical assumption because one of the options for libraries would be to increase budgets, but as will become evident, it seems highly unlikely that the type of budget increases libraries would need in a digital market will ever be feasible. It is easier to visualize than describe (see figure 2.2).

*Figure 2.2*  
E-content and print purchases as a percentage of budget.
would make up about 75 percent of our purchases compared to 25 percent for print materials. It seems safe to assume that most public and academic libraries will experience a similar inversion in the amount of print to digital holdings. The graph is intended to show how radically our holdings would change over a nine-year period in a print-to-digital transition.

Things really get grim when looking at what the cost would do to our overall materials inventory. Given the price of e-books and the five-to-one loss ratio described above, libraries will either have to increase their collection development budget by a five-to-one ratio every five years to maintain the same overall number of items and/or weed less. For public libraries where recently published materials represent the bulk of circulation, more and more funds will need to be diverted to those popular items. The problem, however, is that the loss will be significant enough that it cannot be hidden (see figure 2.3).

The graph visually demonstrates the five-to-one item loss over a nine-year forecast. The most obvious visual is that the e-book collection does not grow fast enough to compensate for the loss of print items. The total collection moves from 250,933 down to 165,719 over nine years. This loss actually is less than it would be if we were to continue weeding at the current rate, but we built in a reduction in weeding a few years into it, because we realized that the shelves would start to look bare or we would need to significantly re-plan the physical space in order to accommodate the loss of items. Ironically, this aggregate loss of items is happening at exactly the same time our patrons are being conditioned to expect more content due to streaming trends. It is possible that libraries will be able to shift music and video budgets into book budgets, since those collections may be irrelevant in this same time period, but such a move to reduce formats goes against the general trends our patrons are experiencing and are conditioned to expect. It is possible to cut databases to pay for e-books, but again, this represents a step backwards toward a significantly less comprehensive collection than before.

For emphasis we put the aggregate loss numbers into a series of pie charts of which the first and last one are presented here (see figure 2.4).

In nine years, we can expect about a 25 percent aggregate loss of materials in our collection. The reduction eventually will slow down when the loss curve and the cost of e-content hit equilibrium, but it is impossible to imagine what will happen to libraries before that happens. In spite of all our efforts to reinvent libraries, our core services still revolve around the lending of materials. Patrons still come to us for content, but if publishers and other content producers finally decide to begin migrating to e-content exclusively, and we find ourselves in an annual licensing model, we will not have enough content to satisfy patrons used to unlimited content supplied by YouTube, Spotify, and the like.

Moreover, this forecast does not even begin to address the fundamental problem between our older “one-user-per-item” model and the streaming model...
that has become the mainstay of the online user experience. While new services like Hoopla are offering streaming content to libraries, a quick survey of Hoopla’s catalog suggests that libraries might be better off paying staff to simply catalog all of the free content available on YouTube, Crackle, and Hulu. Even if such services required constant checking to see when content goes offline, it may still be more cost-effective than paying for yet one more content silo.

In conclusion, we can hope that the above forecast will prove false for a number of reasons. First, the licensing model for e-content has not been settled and may reach a much friendlier price point than the one outlined here. Second, digital first sale may become a reality. But Hollywood and other powerful players are spending millions of dollars working for a model that benefits their bottom line. Libraries can and should be pursuing their current courses of action, but clearly some other lines of action are warranted. It is also possible that the price forecast here may be overly optimistic, in which case the above scenario gets much worse. Forecasting and futurism are a dangerous and imperfect science, especially in the digital economy. This is due not just to the volatility of the radical changes in the foundational economic assumptions like a loss of scarcity, but also because technology continues to evolve and change at a rate no other historic transition can match. While this forecast is serious and significant, unforeseeable changes in the law, the economy, or technology could render it irrelevant in the very near future. Or it could be a very prescient look forward. Regardless, the dangers facing libraries are clear and significant.

Laws can be made through lobbying or they can be written retroactively as a response to an existing situation. Getting out in front of the law the way Douglas County Libraries has done is a critical step, as it demonstrates both market viability and proof of concept to wary content producers. Finally, libraries need to rethink their services and organizations from the ground up. Once content begins making the migration to digital-exclusive formats, everything from our service model to our personnel and hiring and our physical footprint will be challenged. Librarians too often approach these challenges as if they are superficial changes rather than comprehensive ones. Everything we have known will be challenged in a digital world, and we should begin making contingency plans now, because, as we saw in the forecast above, we cannot pretend to survive with 25 percent of our content gone in nine years. We should be planning now for the worst scenarios, and be ready to execute those plans when we see which scenario will eventually play out in the wait for digitization. Above all, we need to advocate and negotiate for an affordable pricing model regardless of what the eventual outcome is.

**About the Author**

Joseph Sanchez is director of Mesa County Public Libraries.