The Espresso Book Machine: A Change Agent for Libraries

Kenning Arlitsch

This is a preprint of an article that originally appeared in the Journal of Library Administration in 2011.


Made available through Montana State University’s ScholarWorks scholarworks.montana.edu
The Espresso Book Machine: A Change Agent for Libraries
By Kenning Arlitsch

Introduction
The debut of the Espresso Book Machine (EBM) in 2006 set off a flurry of speculation among libraries and book publishers about its potential impact, and that speculation has not slowed. Although the list of libraries that own an EBM is still relatively small, the number is growing as On Demand Books continues to expand its catalog offerings, and establishes new relationships that promise to make the machines ubiquitous. This article examines several disruptive implications of the EBM, including user services, collection development, and special collections.

With access to a catalog of millions of digital files the EBM can place a perfect-bound book into a patron’s hands within a few minutes, to borrow and return, to purchase, or perhaps even to keep without paying. In an era of instant gratification, the EBM holds enormous promise for user services.

The EBM has prompted revolutionary discussions about library processes and mission. Collection development practices may be affected, changing a decades-old model from speculative acquisitions to one of acquiring on demand. Some thought leaders envision the library’s future role including “facilities for on-demand printing for consumers, within, or through, the library” (Fraser 2009). Others are considering giving away books, either in print or electronic form, thereby eliminating the costs of cataloging and storage, and in the process completely redefining the library and what it means to have a collection. (Lewis 2010).

Another potential use of the EBM is for printing digitized facsimiles of special collections materials. Paired with speedy scanning technology, such as the Kirtas Book Scanner, it is now becoming much more possible to offer personal copies of original source materials to academics and laypersons alike.

User Services
The website of On Demand Books, LLC, heralds the Espresso Book Machine as revolutionary:

> What Gutenberg’s press did for Europe in the 15th century, digitization and the Espresso Book Machine can do for the world...today! (ODB 2010)

That statement may be audacious, but the EBM does promise to impact print on demand services in libraries. A user who couldn’t find a particular book on the shelf would previously have chosen one of several options: 1) request the book from interlibrary loan and wait several days until it could be delivered; 2) hope the library had access to the digital version of the book in a format that was usable for the reader; 3) travel to another library or bookstore to get a print copy; or 4) give up...
(and many did). But with the EBM the user can request the book from the EspressNet catalog and it will be printed within a few minutes, assuming the book is available in the catalog, and the machine is already warmed up. In this happy scenario the user walks away with a perfect-bound book that is almost indistinguishable from similar books sold in bookstores.

On Demand Books’ website claims the EBM coupled with a high-speed printer will produce a bound 300-page book in less than four minutes. The book size is limited to roughly 8” x 10”, and version 2.0 of the machine can bind up to 830 pages. Costs for the consumables (ink, toner, paper, cover stock) are less than a penny per page. (ODB 2010)

Preparing a “print ready” digital file to the EBM requires some formatting to achieve a consistent and professional level of print quality. (Marriott Library 2010) This is not difficult, but requires standard settings for trim size, margins, tab stops, and spacing, as well as for font types and sizes. Templates can reduce the time and effort needed. The EBM prints a cover, and a good cover design can make the book more attractive and saleable. The content of the book and the cover are submitted to the EBM as separate PDF files.

The chief disadvantages of the EBM are initial cost and complexity. The machine itself costs $85,000 - $100,000, with another $4,000 - $25,000 for the printer, depending on speed and color options. New technology always comes with a high price tag, and over time those costs diminish. Leasing options are also available from ODB, allowing libraries to step into the print on demand world a bit more gingerly. The machines are also complex, with numerous moving parts. They are not yet mass-produced, and their longevity and serviceability may be questionable over the long term.

**Collection Development**

For decades, academic libraries have built their collections by anticipating the needs of researchers, and by creating a set of assets for specific disciplines. Acquisitions budgets supported this speculative purchasing strategy through “use it or lose it” directives. Subject specialists regularly found themselves with end of year funding surpluses that needed spending, and more often than not they found ways to spend down their accounts, but probably not on high-use materials. Academia and the profession itself drove speculative collection building. Libraries that could afford to purchase more were rewarded with larger collections and higher rankings from university accrediting organizations and professional associations.

But studies have shown that the collections themselves have often been lightly used, calling into question the economy of buying on speculation. The well-known University of Pittsburgh survey from the 1970’s was the first significant study that pointed to the dubious returns of speculative collection building. It reported on the use of the Pittsburgh Library collection over a seven year period, demonstrating that “any given book purchased had only slightly better than one chance in two of ever being borrowed.” (Kent 1979) While this assertion was shocking to many librarians
and scholars, subsequent surveys and studies have more or less confirmed the findings. The specific numbers vary by subject and language (O’Neil and Gammon 2009) and by comparison of format (Littman and Connaway 2004), but the premise that print book collections are woefully under-used generally holds true. Despite the data, the debate continues and sometimes takes on an emotional charge, as many librarians revere as archetype the printed book and the large record of knowledge stored in that format.

The Pittsburgh study was published in 1979 but “not much has changed in how academic libraries...select and acquire print titles for their collections” (Branin 2009). More recently, however, the poor economy has made itself felt inside academia, and financial pressures have forced cuts in acquisitions budgets. Seven ARL libraries announced major planned or potential budget cuts in 2009. (DigitalKoans 2009) Even collection-building giants like Harvard University, whose annual acquisitions expenditures from 2004-2008 averaged $29.5 million (Hyman 2009), have found themselves struggling with the realities of the economic recession, and are changing their purchasing strategies. A 2009 Task Force on University Libraries at Harvard urged in its report that the library system “orient its focus more clearly toward ensuring access, as opposed to the current default model of building collections by acquisition.” (Ibid) Ensuring access implies traditional as well as digital formats.

Technological advances are helping to provide options to the traditional acquisitions model. When the Espresso Book Machine was introduced in 2006 it offered a relatively small catalog of books (Rosen 2006). But in 2009, On Demand Books announced an agreement with the Google Books project that expanded the catalog by two million books scanned from major libraries across the country (Single 2009). Combined with original partners Lightning Source and the Open Content Alliance, as well as newly announced agreements (ODB 2010), the EspressNet catalog continues to grow. While the ODB website doesn’t provide a specific number, its EBM brochure claims approximately four million books between the three partners (ODB 2010). Some of those books are undoubtedly duplicates, but the numbers are still impressive and are larger than the print collections of small university and college libraries.

This sudden access to a catalog of millions of books has leveled the playing field, giving smaller libraries that acquire the EBM the immediate access previously only afforded by larger ones. And perhaps even better access. If a library owns a physical book and it is checked out of the collection, the patron who wants it now can’t have it. But an unlimited number of books can theoretically be printed from an electronic file in the EspressNet catalog. This is an obvious and major advantage electronic books have over printed books; many users can read one electronic copy simultaneously, whereas only one reader at a time can use a printed book.

The EBM is only one manifestation of the economic and technological forces exerting pressure on library print collections. The growing number of e-book readers – Kindle, iPad, Sony Reader, Nook, various mobile phone applications – is
another. While some readers still prefer the feel of a paper book in their hands, generational change will undoubtedly see an increase in those who prefer digital renditions. What this pressure means to permanent book collections in academic libraries is unclear, but a gradual reduction in size is almost certain. Some libraries have already begun to take action. In 2005, Murray-Rust recounted a century of shared print repository failures and wondered if they would ever become a reality (Murray-Rust 2005). Five years later Kent State University is planning to reduce its “2.9 million volume collection at a rate of 5 percent per year over the next decade…” (Yonkof 2010) The books will be sent to a shared print repository at OhioLINK. And OCLC will soon release a “Cloud Library” report that will “explore requirements for ‘cloud sourcing’ collections at an academic research library through cooperative agreements with large-scale print and digital repositories.” (Malpas 2009) It is likely that print collections in libraries will begin to shrink, or at least cease growing.

**Lend, Sell, or Give?**
The question of what to do with a book once it’s printed through the EBM is an interesting one. Several options exist. The first inclination of most libraries would be to lend the book and expect its eventual return, at which point it could be added to the existing print collection. That method requires traditional library processing tasks – along with the associated costs – of cataloging and making the book shelf-ready. Some of this process can be delayed until the book has been returned; the immediate requirement is simply to attach a temporary bar code and scan it into the circulation system, thereby keeping the user service as expedient as the EBM makes possible.

Another option is to sell the book. At ODB’s estimates of a penny per page for consumables, a 300-page book from the public domain is a downright bargain. Copyright fees for newer books vary, but even they can often be printed for around $12, even after adding a small fee to cover staffing costs. Selling books moves libraries away from their traditional missions, and can be another cause of consternation among librarians, but the profound effect of disruptive technologies requires a reconsideration of mission. Selling books may provide revenue streams for libraries, or at least help defray the cost of the EBM, and bookstores and copy centers will gladly jump in if libraries do not.

A third option is to simply give the book away. In his recent “thought experiment” Lewis compares the economics of this idea with the traditional model of building a print collection, and concludes that moving entirely to print on demand can work, but it does include risks. License fees could rise and demand could exceed annual budget predictions. These factors would have to be carefully considered, and a hybrid model that requires users to help libraries defray the costs of license fees might have to be implemented. Despite those risks, the argument is compelling. Instead of a pre-purchased print collection that sits on the shelf waiting to be used, printing and giving books away as they are requested would ensure that acquisitions’ spending directly impacts the user. The author concludes that users will inevitably demand that books be available instantly in print or digital form, that
the corpus of print books will be digitized within the next ten years, and that the EBM can play a crucial role in delivering facsimiles of those books. (Lewis 2010). Whether any libraries take on this challenge remains to be seen. This is an area where economic pressures could cause a paralysis of decision. Giving items away for free carries a stigma, even if numbers can be carefully assembled to support the economic viability of the practice.

**Book Scanning at the University of Utah**
The J. Willard Marriott Library at the University of Utah began its digitization program in earnest in early 2000. The first scanning device purchased that year was a Leica S1 Pro digital scanning camera. Mounted on a copy stand, with separate lighting stands and a book cradle, it was the first device the library owned that could appropriately scan fragile books, maps, and three-dimensional objects. It produced spectacular images by virtue of its 25 mega-pixel resolution and Hasselblad lenses. This was an era when we seemed to place a greater emphasis on quality of the scans; general practice was to scan at as high a resolution and bit rate as possible. But each scan averaged three minutes, making digital conversion a tedious and relatively small-scale affair. In the best of scenarios we were limited to producing raw scans at the rate of approximately twenty pages per hour. Still, over the course of several years we scanned hundreds of books, along with maps and art prints.

In late 2009 the library acquired a Kirtas 2400 Book Scanner, and suddenly had the capability of scanning up to 2400 pages per hour. The Kirtas cradle holds a book open at 110 degrees, with two digital SLR cameras mounted on “rabbit ears” above the machine and shooting a frame of each opposing page every time the page is turned. Pages are turned with robotic arms that use light suction to grasp and turn from the center (usually the strongest) section of the page. Once the machine is set up and running, an operator stands by to assure that the pages are turned accurately, and to occasionally flatten a page by hand when the automated brackets fail to do so themselves. The Kirtas scanner also included the BookScan Editor (BSE) software, which provides post-scan image batch processing and optical character recognition to create searchable text for the scanned images. The effect on the scanning rate of special collections materials was dramatic. So too, was the effect on our disk space usage, but that’s an entirely separate conversation.

**EBM Applications at Utah**
The Marriott library purchased an Espresso Book Machine in 2010, with several potential applications in mind:

- Quickly print a book for a patron when the library doesn’t own it, or when it is checked out
- The opportunity for the library to alter its collection development model, as mentioned earlier
- Printing services for local authors who want to self publish their manuscripts
- Replacing commercial printing services for printing student theses and dissertations (Anderson 2010)
Coupling the scanning output of the Kirtas Book Scanner with the EBM to produce facsimiles of unique books held in our special collections

Printing and Selling Digital Facsimiles from Special Collections

Materials held in special collections are often unique and traditionally are used mainly by scholars, graduate students, or serious amateur researchers. When the Marriott Library began digitizing historic newspapers in 2001 we quickly discovered that the lay public – genealogists in particular - were intensely interested in their contents. Historic newspapers had never been very accessible to these researchers. Most of the Utah’s newspapers had been microfilmed as a preservation and space saving measure, beginning in the 1950’s. Trying to find anything in a reel of microfilm was an exercise in frustration, and visually stressful, unless a specific date and page was known. But indexes were almost non-existent. Once the newspapers were digitized with searchable full text and additional metadata, their contents were suddenly unlocked and they experienced more use in a few short years than they had seen in decades. The Utah Digital Newspapers (UDN) remains the library’s most popular digital collection, and is a testament to the effectiveness and importance of digitizing historic collections.

The history of the Mormon migration of the mid-19th century has been preserved in the form of diaries that each of the faithful was implored to keep by Church leaders. These diaries record details of the arduous trip from Nauvoo, Illinois, to Salt Lake City and sometimes beyond. The journeys were often conducted on foot, and with the added burden of pulling a handcart loaded with the family’s possessions. The diaries are rare, many are fragile, and they continue to deteriorate with time. Many Utahns, and Mormons who have moved on to other parts of the country, can trace their lineage to the authors of these diaries. Genealogy is a serious activity in Utah, and while the LDS Church provides extensive resources online, access to these diaries is limited to viewing onsite, or online if they have been digitized.

The Mormon pioneer diaries represent a unique opportunity in Utah, but they are only one example of unique materials in archives and special collections across the country that might find niche commercial markets if facsimiles were offered. Could diaries of Mormon pioneers, or other facsimiles of unique materials be sold? A small ecommerce operation launched by the Marriott Library in 2004 gives some hope that a market may indeed exist. Using the Leica camera described above, as well as flatbed scanners, we made images from a few collections available for purchase as printed posters. We developed rudimentary shopping cart software that integrated with our CONTENTdm display templates, and passed the credit card transactions off to a secure ecommerce site run by the university. Once a credit card was approved, the site transferred order details back to the library, and IT staff fulfilled and shipped the orders. We never marketed or advertised the site, and there was no virtual storefront. The only way to find items to order was from inside the collection database.

Nobody comes to a library to shop. Still, orders trickled in at a small but steady rate, leading us to believe that there was an opportunity, and that what was missing was
appropriate marketing and commercial placement of the digital products. Most libraries do not have staff with these kinds of business skills. One solution might be to take advantage of existing online outlets that could make library products available to shoppers. Online businesses like Amazon, Overstock, and eBay are a few that come to mind. These businesses would take a percentage of sales in exchange for their reach and order fulfillment capabilities. The percentage varies based on how much fulfillment is done by the online outlet.

Some libraries have already ventured into the commercial space, though neither of the following examples focus specifically on special collections. In 2007 Emory University announced that it would begin digitizing its own collections with a Kirtas scanner, and partner with Amazon to fulfill print on demand orders. (Albanese 2007) More recently, the University of Pennsylvania announced a similar partnership with Kirtas Technologies for print on demand, although in that plan nothing would get scanned until a user requested it. (Hadro and Albanese 2009) Again, it remains to be seen whether these efforts will bear fruit, and whether libraries can muster the skills or the willingness to operate in a more commercial arena.

Won’t Google eventually scan Special Collections books?
It’s reasonable to assume that Google may soon enter into agreements with academic libraries and archives to scan their special collections materials. They have more resources and capacity than most libraries, and their progress on the Google Book project has been impressive. Google has scanned approximately fifteen million books in about five years, and estimates that there are approximately 130 million unique books in the world (Google Books). There are currently twenty-one library partners to the Google Books project. At the rate Google is digitizing, and assuming the preliminary Author’s Guild lawsuit settlement is approved (Google Books Settlement Agreement 2009), there will be very few un-digitized books left in ten years.

Not everyone is pleased with Google’s progress, however. Google Book Search (as the public interface of the project is known), has been criticized for the quality of its digitized content. In an article in the Chronicle of Higher Education, Geoffrey Nunberg describes the GBS metadata as “a mishmash wrapped in a muddle wrapped in a mess”, and gives examples of numerous metadata errors he has found in the GBS index. He insists, “extracting metadata mechanically isn’t sufficient for scholarly purposes.” Google’s response has been to clean up the metadata errors Nunberg pointed out, and to begin incorporating library-produced metadata records. (Nunberg 2009)

The quality of Google’s scanned images has also been criticized. Brantley notes “the relatively low quality and lack of uniform image conditions returned by GBS have caused consternation among those seeking to use the GBS repository as a potential source for POD”. (Brantley 2007) He goes on to say that “ineligible scans, missing pages, and image artifacts remain prevalent enough to discourage anything but low-
end reproduction”. (Ibid) Ironically, there are parallels to be drawn with early microfilm processing techniques. The Utah Digital Newspapers program discovered early on that some of the microfilm that was produced in the 1950’s and 1960’s suffered from poor photographic techniques and quality control. Uneven lighting, poor focus, and even hands holding smoldering cigarettes in the image frames were not uncommon, and caused us to seek out print newspapers to scan whenever possible. (Arlitsch and Herbert 2004)

While these statements about Google’s book scanning are discouraging because of the large numbers of books that have already been digitized, they also represent an opportunity for libraries to do it better for their special collections. As Peter Hirtle has noted, the corpus of real one-of-a-kind books is actually relatively small: “Truly unique items will continue to distinguish special collections. Most of the printed items found in special collections, however, are found elsewhere.” (Hirtle 2002) Because the truly unique items are relatively few in numbers, the cost of digitizing them becomes more manageable. Libraries could learn from the mistakes of large-scale digitization efforts and produce a better quality product for their special collections books.

Perhaps in response to these concerns, Kirtas Technologies offers its own digitization services, emphasizing high quality images and post-scan processing. (Kirtas) They include hybrid models where the scanning can be done at Kirtas or at the local institution, with post-processing performed at either site as well. Emory University and the University of Pennsylvania are using variations of these services.

**Where Will Libraries Be?**

Peter Brantley’s prescient 2007 article concluded that print on demand machines “may well occupy the same kind of niche that XEROX machines occupy – an essential component of information delivery that no office/library/school can do without.” (Brantley 2007) In September 2010, On Demand Books and XEROX announced an agreement in which XEROX will “market, sell or lease and service the [EBM] solution, giving retailers new ways to drive revenue. (XEROX). Brantley may have assumed that libraries would be primary targets in the marketing mix of an arrangement like this, but the XEROX announcement makes clear that “retailers”, not libraries, are the targets.

While academic librarians often speak of becoming more agile, the desire is still a long way from the goal. Librarians are not accustomed to making decisions as though our livelihoods depend on them. We are loath to evaluate ourselves based on economic measures, though increasingly we are asked to articulate our value proposition, by our university administrations and by our state legislatures. A recent OCLC study of research library directors found that the “plurality of risks is associated with concerns about an uncertain library value proposition” where the value proposition was defined as a “reduced sense of library relevance from below, above and within.” (Michalko, Malpas and Arcolio 2010) Libraries’ value propositions used to be measured more ephemerally as a “greater good”, but our
funders no longer honor that noble characterization alone. Presenting the value proposition in more measurable terms is a reality we cannot escape.

Libraries are not necessary to the EBM. Any bookstore or copy center with an Internet connection can install one, and many booksellers have already jumped into the fray. An area bookstore that makes the decision to purchase an EBM may fill a service gap that draws traffic to them and away from the library, diminishing our value proposition even further.

Conclusion
The Espresso Book Machine represents an opportunity for libraries to change some long-standing practices and to venture into new areas to better serve users. Several different effects of the EBM and a few potential actions libraries could take have been explored in this article. Library users can derive immediate benefit from a machine that prints books for them in only a few minutes. Second, it is clear that economic and technological pressures are conspiring to change the shape and size of print collections. The EBM can help libraries change their traditional speculative collection development model to one of acquisition on demand, improving economic efficiency and serving users at their point of need. Finally, the EBM, coupled with modern in-house scanning devices or commercial services, can help libraries bring high-quality facsimiles of their unique special collections books to the public, and perhaps even generate a revenue stream that might offset costs.

Libraries are no longer the sole source for researchers, information gatherers, and recreational readers. That reality, along with a significant economic recession, has affected our budgets and has brought increased scrutiny of our practices and value proposition. Libraries have been changing incrementally in response, but we must continually be aware of new concepts that may alter our traditions and missions, even if they seem a bit radical. The EBM is a technological change agent that could help us move in new directions if we are willing.

References


Utah Digital Newspapers [Online] Available at: http://digitalnewspapers.org