The User-Driven Purchase Give Away Library: A Thought Experiment[©]

David W. Lewis
July 2010

Introduction

This article is an exercise in imagination. It is an exploration of a radical alternative to longstanding academic library practice. With this exploration, it is my hope that we will begin to imagine how libraries must change in response what Clay Shirky calls, "the largest increase in expressive capacity in human history."[i] Technology offers new alternatives; in what follows we will explore one alternative way in which these technologies can be applied.

I have made up some of the numbers used in the models described below, but I believe they are sensible and conservative.

Background

For as long as knowledge has been captured in and distributed as documents, libraries have collected documents in anticipation of use. Libraries had no choice as the process of acquiring documents took time and users often could not wait, at least not for several weeks or longer. In addition, documents that went out of print could not be acquired easily or with certainty. Libraries thus bought all of the documents they could afford and organized and made them available to their users. As libraries grew larger they became more complex, more difficult to use, and more expensive. But this was the only viable strategy in a paper-based world where the quality of a library was based on its ability to bring many documents close to the people who might want to use them.

Now, of course, the world is different.

Most information produced today is captured in digital form, though it may be distributed digitally or as a paper document. Increasingly the documents originally produced in paper are being converted to digital formats. The Google Book Project is the largest, but far from the only such effort. What this means is that it is easy to imagine a world where all of recorded knowledge is stored and available in a digital form.

While creating multiple copies and distributing paper documents was hard and expensive, distributing digital documents is fast and easy, and copies are an inevitable result of the distribution process. In the paper world copies of documents were limited in number and they had to be produced in quantity before they were needed and thus warehousing by publishers or in libraries was required. In the digital world as many copies of a document as are needed can be produced at virtually no cost whenever and wherever someone wishes to use it. Warehousing and backup of files is required, but the cost is trivial compared to the comparable costs in the paper world.

Until recently, the changed production, distribution, and economic structures made possible by the conversion from paper to digital technologies could only be partially applied to books because, while transporting and warehousing paper books is cumbersome, the paper book is a remarkably well refined technology that works exceptionally well once it gets into a reader's hands. This has meant that libraries have been understandably reluctant to change their approach to book collections.

Two technologies now becoming available will soon allow the full potential of digital technologies to be applied to books. The first is the development of credible reading machines. While not everyone is enthusiastic many people find that the capabilities offered by various reading machines now provide an acceptable, and sometimes very satisfactory, experience. Books can be acquired quickly wherever you are; dozens of titles can be carried around easily; the font size can be adjusted, and there is a dictionary. These devices are not yet perfect substitutes for paper books, but they are quickly getting better. The second technology is print-on-demand. The ultimate expression of this technology is the Espresso Book Machine. Which creates paper books quickly and cheaply one at a time. Thus we have reached the important tipping point where digital files can be read on machines that are nearly as good as paper books and where paper books can be delivered nearly as quickly, cheaply, and reliably as digital files.

This makes it possible for libraries to radically rethink their fundamental approach to providing documents to users. It is no longer necessary to base the library's strategy on the warehousing of collections.

The Thought Experiment

Let us imagine that we have $1,000,000 a year to spend providing books to a community or an organization.

As a baseline let us imagine a typical library today. Its strategy is to build a book collection. Say the library purchases 10,000 titles at an average purchase price of $35, that the cost to acquire and catalog each of these titles was $25, and that the cost of storing each title for 10 years was also $40. There is remarkably little solid information on such library costs, but these are all reasonable estimates[ii]. Thus the full life cycle cost of the book would be at least $100, and the total cost each year would be $1,000,000. Though some of the cost would be in successive years, the library would have to pay the continuing costs from purchases made in previous years, so it is reasonable to think about a $1,000,000 investment creating a collection that adds 10,000 books per year.

Now let's think about how much use these books might generate. Let's imagine that a quarter of the books generate on average four circulations in their first year in the collection; a quarter generate on average one circulation; a quarter generate on average a half a circulation; and the final quarter generate no use. This collection would thus generate a total of 13,750 circulations in its first year in the collection. Now let's imagine that the number of circulations declines 25% each year. In ten years the total number of circulations generated would be 51,903. This is a cost of $19.27 per circulation. This model accounts for the well-understood 80/20 rule, where 80% of the use comes from 20% of the collection and the equally well understood rule of obsolesce under which book uses decline as a book becomes older.

Now let us consider the radical alternative.

Rather than purchasing books, cataloging them, and putting them on shelves in anticipation of use, what if books were only purchased and produced when a user wished to use it. Rather than loaning the user the book, the library would give it away to the user to keep.

Let's assume that the library leases an Espresso Book Machine for $60,000 a year, pays an operator $40,000 a year, and that on average books printed on the machine cost $3 (a penny a page for 300 pages)[iii]. Further, let us assume that the publisher will sell the rights for a single user version of their texts for $15 (this is half again what Amazon currently charges for Kindle editions), and that this version can be delivered to the single user either as a digital file or as a printed book and that 25% of the users want the digital file. Finally, let us assume that the library has the same $1,000,000 that it would have spent on the 10,000 books it would have acquired if it was functioning in the traditional mode to spend on this user-driven purchase give away strategy.

The question is how many books could the library produce and give away? The answer is 52,174. The average cost to the library to do so would be $19.17. This would include 39,131 paper books and 13,044 digital books. On Demand Books LLC, the producer of the Espresso Book Machine, says a single machine operating twelve hours a day, seven days a week, can produce over 60,000 books per year with minimum supervision, so it seems possible to produce this number of books with one machine. It is important to note that an Espresso Book Machine book is not as high quality a book as the hardbound volume typically added to library collections, but it is not bad.

It might make sense for the library to charge a modest fee for this service to discourage users from overly casual use, say $3, or the marginal cost, for a printed book and $1 for a digital book. If the $152,542 earned from these charges is reinvested in the service, 61,017 books could be produced, 45,763 paper books and 15,254 digital books.

In the first case the $1,000,000 investment would create about the same amount of use as the traditional collection-based strategy. With modest fees, it does better.

Importantly for users the books they were looking for would always be available. Users would never find their ability to use a book stymied because someone else had it checked out. In addition, since the user would permanently own the book, one might argue that it would be more valuable to the user than a book borrowed form the library. They could underline, highlight, or write in the margins.

Discussion

The scenario for a User-Driven Purchase Give Away Library presented in the thought experiment above suggests that the library could operate in a fundamentally different way.

This possibility raises several interesting issues.

Don't We Want the Books Back?

The short answer is "no". If the book is returned to the library it has the potential to be used in the future, but it imposes the significant costs of cataloging the book and storing it on the library. It is also likely to impact the cost of the license publishers would charge, at least for digital files.

To demonstrate this, let us assume that rather than giving the books produced in the user-driven purchase arrangement away, we rather loan them to the user and when they are returned the book is cataloged and kept for later use.

Thus there would be a cost of cataloging, say $15, and the cost of storing the book for ten years, or $40. The book then has a cost, not of $3 but of $58. If we assume we keep the $1 fee for the digital books, which we assume will not be returned and have no fee for the paper books that will be returned, then our $1,000,000 investment would produce 14,230 books, 10,673 paper books and 3,558 digital books.

The circulation pattern for a collection of books acquired with a user-driven purchase plan can be expected to be different from a traditionally selected collection because one of the few things we know about book use is that the only good predictor of future use is past use. If we assume that over ten years the most used quarter of the paper books will be used five times (probably the useful life of an Espresso Book Machine book); the second most used quarter of the books will be used three times; the third most used quarter will be used two times; and the least used quarter only one time, then the total number of uses would be 29,349 paper book uses. Add in the 3,558 digital book uses and the total uses would be 32,907. This is less use than the traditional library model produces and is below the number of uses produced when the books are given away.

Pursuing only half the strategy gets a poorer result than not pursuing it at all.

Not Cheap Enough

While the User-Driven Purchase Give Away Library creates an arguably better investment than the traditional library, the model presented above would likely fail because it will be easier and quicker to use, and because nearly any desired book would be available. It would be easy to imagine that demand might be two or three times what the model above proposes as possible.

One down side of the User-Driven Purchase Give Away Library is that once the promise to provide books has been made and fees have been set, the library has an obligation to deliver. Suspending the service midyear because you have miscalculated the cost or, more likely, demand and have thus exhausted the library's budget would not be a particularly pleasant outcome. This is a type of risk that is absent in traditional library practice. The User-Driven Purchase Give Away Library requires that we be prepared for "catastrophic success".

With this in mind it would be interesting to explore what would be required to meet demand at higher levels. As it turns out there are really only two constraints, the first is the capacity of the Espresso Book Machine. If you print

more than 40,000 to 50,000 books, you will likely need to lease and staff a second machine. The second constraint is the cost of the rights. Or, more precisely, the amount the library is subsidizing this cost.

Two things could be done to reduce the amount the library spends to procure these rights. First it could negotiate a lower price. Second, it could ask users to pay a larger portion of the rights costs. The first strategy, especially if pursued collectively by libraries from multiple institutions, might be productive.

The negotiation of a site license where cost would be independent of use would be the best solution. This is a common approach for many vendors who provide libraries other electronic products, so it is not unimaginable.   Certain books, bestsellers, for example, might never be available in this mode, but it is easy to imagine collections of university press titles being offered in this way. If the Google Book deal ever gets done, this type of licensing seems likely.

Finally, it is not unreasonable to expect more public domain and open access scholarly monographs to become available. The library's ability to make these digital files into a printed book of reasonable quality and at a reasonable price then becomes a valuable service.

Discovery

One persistent claim for a physical collection of books is that it aids in discovery. That is the serendipitous discovery of an unexpected useful book because it is near the book you were originally looking for. Browsing in a physical collection clearly has some value, even though its limitations are often understated and its virtues romanticized. The limitations are clear. A book in a physical collection can be in only one place, even though it can share intellectual connections with many other books in many other places. Many of the best books are in circulation and therefore can't be found by browsing. Finally, the library may not have acquired a useful book.

Digital systems offer alternatives to browsing. Amazon has a variety of recommendation systems and Google Book Search provides full-text search across millions of titles. WorldCat provides a more traditional library catalog approach to discovery for the aggregated collections of thousands of libraries, and Library Thing uses a social networking strategy.

Discovery and filtering of results will be key to successful use of the vast array of web-based resources that will only continue to grow. While it is probably unwise to try to predict what strategies and tools will emerge, it is inevitable that effective tools will be found. The end of browsing will be lamented by some, but at the end of the day, different and better strategies to aid discovery will be found.

<u>The Library as an Information Subsidy</u>

I would argue that at its core the User-Driven Purchase Give Away Library is serving the same purpose that traditional libraries have always served. I would argue that at its core either type of library is the means that communities and organizations use to provide a subsidy for information use. Communities and organizations provide this subsidy because they understand that left to their own devices individuals will not pay for as much information as they should to be fully contributing members and citizens.[iv]

In a traditional library the subsidy is used to provide for the building of a collection of books and organizing them. Use of the collection is generally free, at least to members of the community or organization, though it often requires a non-trivial investment of time on the user's part. Traditionally large book collections have been viewed as an institutional or community asset, though the long-term commitment to a book collection also creates a significant liability.

In the past, the possible future use justified the liability of a physical book collection. In the easily imaginable future, alternative delivery mechanisms call this justification into question.

In the User-Driven Purchase Give Away Library the subsidy is provided by funding the print-on-demand capability and in offsetting the cost of acquiring the single user rights to books. In the modest fee model the user pays the incremental cost of producing the paper book and contributes to the cost of the rights. Even with such a fee, it is important to note that the cost to the user may be significantly reduced. This is because the time required to use the User-Driven Purchase Give Away Library would be much less than a traditional library. A digital file could be delivered instantly to anywhere and paper books could be ordered from anywhere and picked up when it was convenient for the user or delivered to the user's home or office. But of more importance is the fact that the User-Driven Purchase Give Away Library can potentially have all of the books, whereas the traditional library has only a subset of what exists. The importance of this is that the uncertainty of finding the book you want that users experience in a traditional library, either because the library did not purchase it or because another user has it checked out, does not exist in a User-Driven Purchase Give Away Library. This uncertainty is a major contributor to the cost of using a traditional library.[v]

Conclusion

It is hard to imagine, even if all the pieces required to make the User-Driven Purchase Give Away Library possible were in place, that any library would

embrace it today. But if we imagine what the world is likely to be like in ten years, I would suggest, it can begin to seem downright sensible.

Ten years from now it is likely that the historic corpus of printed books will have been converted to digital files. Google will hold them, but so will the HathiTrust and other academic or not-for-profit organizations. In addition, print copies will be stored in long-term print repositories. Both digital and physical copies will be available and their long-term preservation will be secured. This will mean that the space required for traditional local library collections will become much less necessary. The pressure to repurpose this space will quickly become irresistible.

Ten years from now digital book readers will be common and print-on-demand machines will be better and cheaper.

Ten years from now many publishers will have been forced to find new economic models as the number of print copies of books purchased continues to fall. This is likely to include open access models and site licenses.

Finally, ten years from now everyone will expect that all documents should be instantly available anywhere in all of the forms — digital or paper — that might be useful.

In this future, some money will still have to change hands, but much cost can and will be wrung out of the system. Among the cost that should be eliminated are the costs of selecting, organizing, and storing local library collections. The reasonable costs that will remain should be in the editorial activities that add value to manuscripts. Users will pay some of these costs, but some will continue to be managed as organizational or community subsidies.

In light of these changes, it seems to me that the User-Driven Purchase Give Away Library may not only be likely, but it could be inevitable.

Notes

[i] Clay Shirky, "How Social Media Can Make History," TED Talk, June 2009. Available at:
http://www.ted.com/talks/clay_shirky_how_cellphones_twitter_facebook_can_make_history.html

[ii] In 1997/98 the cataloging cost at Iowa State University was estimated at $16.25. See: Dilys E. Morris, Collin B. Hobert; Lori Osmus, and Gregory Wool, "Cataloging Staff Costs Revisited," *Library Resources & Technical Services* 44(2):70-83 April 2000. Paul N. Courant and Matthew "Buzzy" Nielsen have recently estimated the cost of storing a book in open stacks at $4.26 per year with a discounted life cycle storage cost of $141.89. See: Paul N. Courant and Matthew "Buzzy" Nielsen, "On the Cost of Keeping a Book," in *The Idea of Order: Transforming Research Collections for 21st Century Scholarship*, Washington, DC: Council on Library and Information Resources, June 2010. Pages 81-105. See in particular the chart on page 91. Available at: http://www.clir.org/pubs/abstract/pub147abst.html. Stephen R. Lawrence, Lynn Silipigni Connaway, and Keith H. Brigham estimate the discounted life cycle costs on acquiring, cataloging and storing a monograph at $119.56 based on 1999 ARL data. This is 265% of the purchase price of the monograph itself. See: Stephen R. Lawrence, Lynn Silipigni Connaway, and Keith H. Brigham, "Life Cycle Costs of Library Collections: Creation of Effective Performance and Cost Metrics for Library Resources," College & Research Libraries, 62(6):541-553 November 2001. See in particular Appendix B on page 553. Available at: http://crl.acrl.org/content/62/6/541.short.

[iii] See the information provided on the On Demand Books, LLC website at: http://www.ondemandbooks.com/home.htm

[iv] See Lewis, David W. "What If Libraries Are Artifact Bound Institutions?" *Information Technology and Libraries* 17(4):191-197 December 1998. Available at: http://idea.iupui.edu/dspace/handle/1805/434.

[v] For an analysis of the effects on uncertainty on the cost of using a traditional library see: Lewis, David W. "Why Books are Bought and Borrowed." *Bottom Line: A Financial Magazine for Librarians* 2(4):21-24 December 1988. Available at: http://idea.iupui.edu/handle/1805/1170