CAN'T BUY US LOVE

(Rick Anderson)

45 R.P.M
4:23

Series 001
082013
(S&R 32582)

ITHAKA S+R

CAN'T BUY US LOVE

The Declining Importance of Library Books and the Rising Importance of Special Collections

RICK ANDERSON

Interim Dean & University Librarian, J. Willard Marriott Library, University of Utah

Introduction

Research libraries throughout North America are experiencing a massive decline in the use of their general collections[1]—their large and comprehensive collections of printed books and journal volumes purchased in the commercial marketplace. This decline is the inevitable outcome of a massive shift in scholarly publishing from an analog and print–based model to a digital and networked one. In this environment, it is no longer obvious that it makes sense for research libraries to continue their traditional practice of creating, housing, and maintaining such collections. In this paper, I will propose that we shift our focus from the collection of what we might call "commodity" documents (especially in physical formats) to another activity in which we have also been engaging for many centuries: the gathering and curating of rare and unique documents, including primary–source materials.

The radical shift is not the format shift

For centuries, readers and researchers have relied on academic libraries to provide them with access to books that are sold (or access to which is sold) in the commercial marketplace. Patrons relied on libraries to solve what were essentially problems of market inefficiency: books were expensive physical objects that cost a lot to create, move, organize, and maintain. Since books were expensive, no one could buy all the books he or she needed, and since they were bulky and heavy and fragile and prone to loss, the cost of housing and caring for large collections of books was prohibitive even if the cost of buying them was not. By pooling institutional resources, research libraries made access to these expensive resources available, on a shared basis, to all in their host institutions, and undertook their housing and preservation as well.

The same thing has remained true even after scholarly communication moved largely out of the physical and analog environment and into a networked and digital environment: access to online journals (which are used far more heavily than monographs, especially printed monographs, in research libraries) and ebooks remains very expensive, even though the costs of processing and "housing" online documents is dramatically lower. This has meant that throughout much of the 1990s and early 2000s the role of the research library as a broker (buying access on behalf of a large community of users), a curator (ensuring that access endures), and an organizer (making resources easier to find and use) remained important even as the acquisition and housing of research content moved substantially into a virtual environment.

However, during the past two decades the networked digital environment has done more than just become the default locus of scholarly communication. It has done three other things as well.

1. It has made the market for "commodity" books much more efficient, driving the price of books down. By "price," I do not mean average list price—which has remained relatively stable over time[2]—but rather the actual amount of money typically paid for a copy of a book. In the past, if one wanted to buy a book, the cheapest option available was the cheapest local option: if the lowest price in a nearby bookstore was $20, that was the best deal most buyers could hope for. With the advent of such internet–based outlets as Amazon Marketplace and Bookfinder.com, however, every home with an internet connection has direct access to the holdings of thousands and thousands of bookstores around the world, and the likelihood of finding a remaindered or used copy—often at a price of literally pennies, plus a few dollars in shipping—is very high. Prior to 1995, a library had to be careful not to lose its 1975 printing of, say, Steinbeck's *East of Eden*, and had to take good care of that copy lest it fall to pieces from decay or use. In 2013, the loss of the library's physical copy of a commodity book has little if any impact on its patrons' access to that book. As of this writing, replacement copies of *East of Eden* can be had from more than a hundred online

[1] Rick Anderson, "Print on the Margins: Circulation Trends in Major Research Libraries," Library Journal 136 (2011): 38–39.

[2] Brent Cox, "How Much More Do Books Cost Today?," The Awl (2011), accessed June 14, 2013, http://www.theawl.com/2011/12/how–much–more–do–books–cost–today

booksellers at a cost of one dollar or less, plus shipping. This means not only that the library can afford to take less scrupulous care of its existing copy, but also—and much more importantly—that the library's patrons simply no longer depend on the library for access to that book in the way they once did.

2. The new information environment has greatly reduced the cost of information distribution, making it much easier for scholars to share documents (such as articles that they or others have written) with each other. New technology has also made it possible to digitize and make publicly available an enormous corpus of public–domain books—physical copies of which still cost money in the commercial marketplace, and which of course also cost money to house and care for. Online aggregations like Google Books and HathiTrust, by putting digital copies of public–domain titles into a publicly available online database, have effectively obviated the need for most libraries to hold physical copies of those books. By joining HathiTrust in 2010 and thus gaining full–text access (along with full download rights) to that collection's 3.3 million public–domain books, the research library in which I work effectively doubled the size of its book collection—and at a trivial cost. Costs of distribution should not be confused with costs of publication; however, the ease with which documents may now be distributed after (or before) publication means that scholars no longer rely on libraries for access to those documents in the way they once did. Professors whose libraries do not subscribe to the journal containing a needed article would once have argued strenuously for a subscription or requested a copy via interlibrary loan; today, they are just as likely to contact the article's author directly and request a copy by email, in many cases receiving a positive response within hours or even minutes. This is admittedly not a terribly efficient way to get access to journal articles, and the pressure on libraries to maintain strong subscription lists remains great. But the fact that such an informal acquisition network now exists leads to the third impact that a pervasively networked digital information environment has had on the scholarly communication system:

3. It has made possible the Open Access (OA) movement, which could never have existed in a print–based information environment. The movement's success to this point has relied, in part, on the widely held perception that the digital environment has virtually eliminated publication costs, in part on assertions about what the public pays for when it funds scientific research, and in part on the undeniable fact that once a document has been placed on the network, the cost of creating and distributing an additional copy of that document falls to virtually nothing. The degree to which institutional or governmental OA policies, combined with publishers' adoption of author–pays OA models, will succeed at turning what were once commodity documents into effectively public property remains to be seen, but the movement's success has been considerable so far and it continues to grow. (Although, at this point, it applies almost entirely to the scholarly and scientific journal marketplace and not to books.)

What makes these three developments significant for the future of libraries is that, together, they point to a single, potentially transformative reality.

Focusing on the wrong shift

When, in the 1990s, it became clear that the scholarly communication system was moving almost entirely out of the print realm and onto a worldwide digital network, we in libraries immediately began dealing with this development as a format shift: whereas we had been brokers, curators, and organizers of print collections, we now prepared ourselves to become brokers, curators, and organizers of online collections. We focused on what this shift would mean for our workflows, for patron access, for privacy, for preservation,
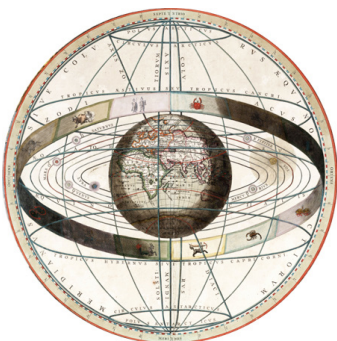
and for our traditional discovery tools. Much ink was spilled and many trees killed as we argued about how best to address and adjust to these changes—but in general, I believe we have made that adjustment quite well. The typical research library does an admirable job of making enormous amounts of relevant and high-quality information available with reasonable ease to its patrons in a variety of formats, and has created a new superstructure of services designed to ensure reasonable permanence of access where such is needed. We have established buying consortia to give our budgets extra leverage and to create economies of scale, and have built tools to make management of online resources more effective and more cost-efficient. Our response to the print-to-online shift has not been perfect, but we have generally been successful at turning ourselves from brokers, curators, and organizers of print materials into brokers, curators, and organizers of online digital materials.

However, during these past two decades of radical change and energetic response, I believe we have missed a much more important shift, one that poses a more direct and existential threat than the one posed by the move from print to online. We have failed to prepare for the emergence of a reality in which our very role as brokers, curators, and organizers has itself been fundamentally undermined.

In other words, the gap that should most concern us in research libraries today is not the one that lies between *physical and online documents*, but the one that lies between *commodity and non-commodity documents*. The library's important role as a broker arose largely because of the many inherent inefficiencies of a print-based information marketplace. Today's more efficient online marketplace features much lower prices and much lower barriers to personal collection-building, a pervasive full-text searching capability that makes traditional cataloging less obviously necessary, and widely distributed storage and access points that undermine traditional approaches to preservation and curation. This new reality affects commodity documents no matter the format, though its impact is most obvious in the book marketplace—particularly the market for trade books that are produced in large print runs. (It is less the case in the journal economy, where there is no appreciable secondary market, and therefore quickly rising subscription prices pose a much higher barrier to individual access.) The bottom line is that libraries are no longer needed in the way they once were to provide access to documents that are available in the commercial marketplace.

The Janus-faced library

If the library's role as a broker, curator, and organizer of commodity documents is fading, what significant roles remain?

Virtually every academic library is really two very different organizations housed in a single building and united by an artificial administrative connection. The first library, the one everyone sees, is that which manages a collection of commodity documents, in both print and online formats. The library imposes less control on this collection, allowing its contents to be used without supervision; this is because the materials in the general collection tend not to be unique or even rare, and can usually be repaired or replaced without any loss of value to the patron. The second is the library that we typically call Special Collections. Materials in Special Collections are typically much more tightly controlled, because they are not easily replaced and physical repair might significantly reduce their value. A library that acquires a copy of Andreas Cellarius' 17th-century *Harmonia Macrocosmica* is not doing so because that document will provide the astronomy faculty good information about the solar system, or even to show what 17th-century scientists believed about the solar system (which could be accomplished much more cheaply and conveniently with a transcription or a facsimile copy). It is doing so in order to preserve and provide carefully controlled access to a beautiful and rare document the

value of which has as much to do with illuminating the history of binding and printing as the history of science.

Thus, whereas the main library buys documents primarily because of their *curricular relevance and instrumental value*, Special Collections is often more interested in documents as *artifacts that are valuable in and of themselves*, independent of the immediate relevance or utility of the words and ideas they contain—documents which, in a great many cases, are rare or fragile enough to be in serious danger of loss without special care. Thus, a mass-market 1975 printing of *East of Eden* will be shelved in a public area of the library, where patrons may take it down and peruse it at will and even take it home with the promise eventually to return it; if the book falls apart, it will likely not be repaired, but rather replaced—perhaps even with a different edition. In the circulating collection, the content matters much more than the container. Special Collections, on the other hand, is more likely to hold a signed first edition of that same book, which will be made available for use by patrons only under supervision and in a tightly controlled environment. Under no circumstances will a patron be allowed to remove that item from the library. If it is damaged, replacement would be extremely difficult and repair costly; in the case of unique materials, replacement would be simply impossible. In Special Collections the content matters, but the container often matters more.

The rare and unique materials held by Special Collections are, in other words, not commodity documents. While they may be bought and sold in a marketplace, it is a highly specialized marketplace with a unique structure and unique rules. It is not Amazon, and most members of the general public do not have ready access to it. In quite a few cases, in fact, the materials held by a research library's Special Collections are not purchased by the library at all, but are given to the library by donors working outside the commercial marketplace entirely.

The distinction that will shape our future

It is the opportunities inherent in this distinction—the distinction between commodity documents and non-commodity documents—that I believe will bear heavily on the future utility and health of the academic research library. The print/online dichotomy is no longer a terribly meaningful one, and indeed it may have been a red herring from the start, mere camouflage for the real shift that was happening, which was a dramatic increase in the efficiency of the marketplace for commodity documents. As the market in those documents continues to grow in efficiency, and as unassisted discovery of those documents continues to become easier and easier, many of the academic library's traditional roles are moving to the margins of the research experience. In short, very few academic patrons truly rely on their library to buy, process, describe, and preserve a 1975 printing of *East of Eden*. However, if that library's Special Collections owns a handwritten poem by a 9th-century Arab poet, or a handwritten journal from a Civil War combattant, and fails to curate that item, the entire world (not just the library's local constituency) suffers—particularly in situations where the library is holding a collection on behalf of a community which lacks a strong archiving or publishing infrastructure of its own.

What, then, does curating such a document entail? In the past, it meant keeping the document in a tightly controlled protective environment and drastically restricting access to it. In today's environment, it means doing the same—but much, much more as well.

What the world needs research libraries to do now—and this need is both powerful and growing—is provide broad and easy access to the intellectual content of rare and unique non-commodity documents that would otherwise remain unfindable and unusable. This means:

1. *Acquiring them.* Rare and unique documents held by private collectors, or languishing unknown in archives and basements, do the world of scholarship very little good. Libraries have money to spend on documents, and need to direct more of that money away from the purchase of commodity documents that are already relatively freely and easily available to the public and toward the acquisition of documents that, unless they are held by libraries, will not be made available to scholars in any meaningful way.

2. *Digitizing them.* A library that owns a rare or unique book and simply keeps it locked up is doing the world no greater service than a private collector who does the same. While physical access will generally have to be restricted in order to preserve the document, access to the *words, ideas, and images* contained in the document can and should be made available as widely as possible. This can be done most effectively by creating high-resolution images of those contents and disseminating the images online.

3. *Making them discoverable.* Access is not only limited by format, policy, and practice; it is also limited by findability. What makes documents findable is good metadata (including, as needed, transcription and/or translation), organized and optimized to expose itself promiscuously to popular search engines. A document that cannot be found is, for all intents and purposes, nonexistent.

Uniqueness and mission

There is a potentially serious barrier for any library that wishes to move aggressively in the direction I have described: that barrier is the mission of its host institution. For a major research library to suddenly and unilaterally redirect, say, 50% of its collections budget to the acquisition and processing of rare and unique materials would be both irresponsible and unwise. We must bear in mind that while such materials will, in many cases, offer tremendous value to the wider world of scholarship, and while in some cases it may be possible (and even tremendously beneficial) to incorporate those materials into the local curriculum, rare and unique documents may not always provide direct support to the most centrally important mission elements and strategic directions of any individual library's host institution.

This implies several important realities.

1. Every research library must *strike the right balance* between benefiting the larger world of scholarship and supporting its institution's specific teaching and learning and research goals. If it fails to do so, it will lose the support of its host—and rightly so. Furthermore, what constitutes the right balance should (and, in the long run, will) be determined by the host institution, not by the library.

2. Most of the people served by a research library spend little or no time thinking about the difference between commodity and non-commodity documents, and may be only peripherally aware of the tremendous changes that have taken place in the scholarly communication economy. In order to gain support for some degree of shift in priorities away from traditional commodity documents and towards non-commodity ones, librarians will have to explain *clearly, concisely, and compellingly* why such a shift makes sense and how it will be beneficial in terms of both local and broader public good. Support for the shift will be determined by a combination of high-level institutional buy-in and general acceptance of the explanation by the library's on-the-ground patron constituency.

3. For most libraries, the shift will have to be gradual. Key to success will be enacting the

shift in a *steady, realistic, and wholly transparent* manner. The appropriate speed and trajectory of the shift will vary from library to library, and will have to be determined in consultation with institutional administrators. At all points, it will have to be enacted in an open and transparent way.

The qualifying language in each of the above points is absolutely essential. It is a hard truth, but true nevertheless, that an academic library does not (and should not) define its own role on campus. It can (and should) actively contribute to the shaping of its role, and library leaders absolutely must both develop and promulgate an expansive and ambitious vision. But veto power lies with the leadership of the host institution. The academic library exists to move the university forward, not vice versa. Library leaders who lose sight of this fundamental fact will eventually lose their jobs, and rightly so.

Opting out of the scholarly communication wars

There is an ancillary but significant benefit to shifting the focus from commodity to non-commodity documents. The politics and economics of scholarly communication are increasingly fraught. Publishers (both for-profit and not-for-profit) working in the commercial marketplace defend their revenue streams; libraries fight to keep prices down; authors submit their manuscripts to the publications most likely to help them secure prestige and tenure while (for the most part) trying to stay out of the skirmishes constantly breaking out between publishers and libraries; and readers either get access to scholarly publications or do not, depending on their institutional affiliations or their personal buying power.

Over the past two decades, librarians have grown increasingly frustrated with the existing scholarly communication system, calling on each other to create an alternative, noncommercial structure based in the academy[3] rather than the marketplace, and calling on authors to stop supporting the old one. But the traditional system exerts a very powerful gravity—so powerful, in fact, that even when OA models have taken hold and begun to flourish, they have generally been absorbed into the traditional system rather than subverting it. The so-called Gold OA model[4] (the one that has been willingly adopted by many STM publishers and functionally enshrined in national policy thanks to the UK's Finch Report[5]) does not undermine the old system so much as enrich it, and while it does result in better access for readers, it also shifts costs to funding agencies and research budgets—and thus away from the support of research itself. Whether one sees this as a good thing or a bad thing on balance depends on one's perspective and goals. Those who were hoping that OA would put commercial science publishers out of business can only be disappointed by such developments, whereas those for whom expanded access is most important have reason to be happy about them.

[3] "LPC Project Background," Educopia Institute, accessed June 18, 2013, http://www.educopia.org/programs/lpc.

[4] Wikipedia contributors, "Open Access," Wikipedia, the Free Encyclopedia, accessed June 14, 2013, http://en.wikipedia.org/wiki/Open_access.

[5] "Finch Report," Research Information Network, accessed June 14, 2013, http://www.researchinfonet.org/publish/finch.

But a shift in focus from the brokerage of commodity documents to the gathering, processing, management, and wide distribution of non-commodity documents allows us to sidestep the whole Open-Access-versus-toll-access controversy. Non-commodity documents are, by definition, not for sale—or at least not in the same way that commodity documents are. A library that shifts a portion of its budget and staff time in the direction of making noncommercial documents more findable and accessible is neither undermining the existing scholarly communication system (except to the extent that it pulls collections money away from commercial purchases) nor supporting it. Instead, it is contributing to a separate system, *one that feeds the scholarly conversation without exerting control over it*—the non-commodity information environment supports scholars without either restricting access to the documents (as the toll-access commodity model does) or constraining authors' publication options (as OA mandates do).

The situation in one large research library
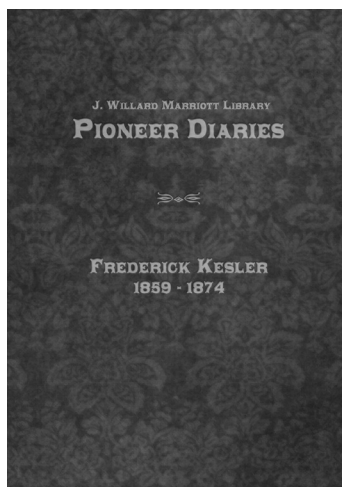
Allow me to close with an illustration.

The research library in which I work, at the University of Utah, holds in its special collections a number of handwritten diaries produced by 19th-century pioneers who came west on the Overland Trail. These documents are unique and fragile, and therefore not suitable for circulation or lending, and physical access to them can be granted to researchers only under tightly controlled conditions. They are handwritten in often hard-to-read script, and therefore need transcription in order to make their texts machine-readable and electronically searchable. In addition to full-text transcription, making them discoverable by researchers will also require accurate and reasonably comprehensive metadata, including reliable authority records that will ensure they are grouped with similar or related items in search results.

We do not know exactly how many of these diaries we have in our special collections, because they are distributed among multiple donations and subcollections that have yet to be fully registered at even the box level, let alone cataloged at the item level. Given the relatively small staff allocated to Special Collections, it will likely be many years before these diaries are all found and processed—and by that point, many other rare and unique documents will have joined the queue. These diaries represent an incredibly rich and potentially useful set of research materials, and they are only a tiny drop in the enormous bucket of rare and unique documents our library owns. Most of these are virtually unfindable and effectively unusable, because the focus of our library has always been, like that of most libraries, on the gathering, organization, and management of commodity documents. Furthermore, our library is, in all of these ways, typical of large research libraries throughout the United States—which means that our special collections in their entirety represent only one drop in the much larger bucket of special collections that currently languish, undigitized, undescribed, unfindable, and inaccessible, in similar libraries throughout our country and the world.

Our library's main, public collection, on the other hand, includes several million printed commodity books. None of these books is unique and very few are even rare; if lost or damaged, they can virtually all be replaced, generally quite cheaply. A significant portion of them now exists in online versions that are freely available to the public[6], both for online reading and for downloading in full text. Metadata has already been assigned to these books, and in the vast majority of cases had already been assigned to them even before we purchased and processed them ourselves. A small and fast-shrinking number of these books is checked out or even consulted by students and faculty in any given year, and yet their acquisition and management absorbs roughly 25% of our library's total fund of staff time and they occupy a similar share of the library's increasingly crowded floor space.

Pioneer Diary example

Directions for the (near) future

Does access to commodity documents matter in a research library? Of course it does. I want to be very clear that I am not advocating that research libraries abandon the brokerage and management of these documents. I am, however, suggesting that research libraries devote a greater percentage of budget and staff time than we hitherto have to the management and dissemination of those rare and unique documents that each of us owns, that no one but the holder can make available to the world, that have the potential greatly to enrich the world of scholarship, and that can be made available outside of the commercial marketplace without damage to any participant in the scholarly communication system. Importantly, I am also urging that each of us make this shift in consultation with our local stakeholders and in harmony with the missions of our host institutions.

One final point: as we begin to move in this direction, it is imperative that we avoid confusing uniqueness with value. The goal of the shift I am describing is not to make the

[6] "HathiTrust Dates – Public Domain," HathiTrust, accessed June 14, 2013, http://www.hathitrust.org/visualizations_dates_pd.

library or its offerings "more unique." Uniqueness may be an important characteristic of a valuable collection, but it has little bearing on value in and of itself. (Given a pen and paper, any one of us could create a perfectly unique but utterly worthless document at a moment's notice.) The goal is to enrich the scholarly environment with useful books and other documents that would otherwise remain hidden from scholars and students, and to shift our focus from resources and activities that make relatively little contribution to that environment to those that will have the greatest enriching effect upon it.