From Stacks to the Web: The Transformation of Academic Library Collecting

David W. Lewis

The existence of a ubiquitous and cheap worldwide communications network that increasingly makes documents easily and freely available will require a transformation of academic library collecting practice. It will be driven by a number of specific developments including: the digitization of content; the development of print repositories; the development of e-readers and print-on-demand publishing; the growth of open access; challenges to establish academic publishing organizations; and the growth of new forms of scholarship based on openness and social productivity. If academic libraries are to be successful, they will need to: deconstruct legacy print collections; move from item-by-item book selection to purchase-on-demand and subscriptions; manage the transition to open access journals; focus on curating unique items; and develop new mechanisms for funding national infrastructure.

Different parts of the Ocean contained different sorts of stories, and as all the stories that had ever been told and many that were still in the process of being invented could be found here, the Ocean of the Streams of Story was in fact the biggest library in the universe. And because the stories were held here in fluid form, they retained the ability to change, to become new versions of themselves, to join up with other stories and so become yet other stories; so that unlike a library of books the Ocean of the Streams of Story was much more than a storeroom of yarns. It was not dead but alive.

“Salman Rushdie, Haroun and the Sea of Stories”

Academic libraries are about to undergo a transformation in their collection practices. The change will be fundamental and will affect nearly every aspect of library service. It will require a rethinking of budgets and alter the longstanding collaborative relationships between academic libraries. It will require an examination of the fundamental practices and values of academic librarianship. Once we make sense of what we need to do and how we need to do it, we will have to explain it to faculty, many of whom will be appalled and angry, and administrators, many of whom will hear only the parts of the mes-

David W. Lewis is Dean of the University Library, Indiana University Purdue University Indianapolis; e-mail: dlewis@iupui.edu. © 2013 David W. Lewis, Attribution-NonCommercial (http://creativecommons.org/licenses/by-nc/3.0/) CC BY-NC
sage that might lead to budget reductions. All of this will make for interesting times.

There are specific drivers of the transformation that will be discussed in some detail below, but the fundamental force for change is the growth of computing power and its application to create a ubiquitous and cheap worldwide communications network. This network contains a large and growing amount of content. It is beginning to resemble Rushdie's Ocean of the Streams of Story.

Computing technology has been impacting library practice for nearly a half century, and it is tempting to think that the pace of change will continue more or less as it has up to now. But this is not what will happen. We have reached a tipping point; the pace of change is about to accelerate. The book is departing its physical codex manifestation and will quickly join most other documents as digital entities on the web. Libraries will soon be able to provide much of the information required by their communities without resorting to local print collections. Print collections have required a specific set of practices. These practices have been developed and refined over hundreds of years and are reinforced with a set of values that define what excellence is for a library and librarians. Digital documents will require a very different set of practices and, at least in some cases, new values. The impact of these changes will be felt not only by libraries but also by all of the other players in the scholarly communications chain: authors, publishers, book and journal distributors, bookstores, and promotion and tenure committees.

Michael Buckland: The Issues Defined and the Future Foretold
Twenty years ago, Michael Buckland published Redesigning Library Services: A Manifesto. In this short monograph, Buckland provides a structure for viewing library history and parses the core aspects of library practice. He also looks forward and provides key insights for how to think about the world we are now enter-
several copies be stored, “carefully at
different locations under suitable condi-
tions.” The specifics will be different for
paper and electronic documents, he says,
but the general pattern will be similar
for both. In the dispensing role, digital
documents will lead to “a much reduced
premium on local storage compared with
storage at a distance.”

As Buckland considers the changes in
collection practice that digital documents
would bring, he points out that the dis-
pensing function accounts for “the great
preponderance of operating costs and
space needs in the Paper Library and in
the Automated Library.” Because the
demand for library materials is unevenly
distributed and difficult to predict, most
of the cost of collections is for large quan-
tities of relatively little-used material.
Buckland also reflects on the symbolic
role of library collections, how large paper
collections provide institutions with sta-
tus and prestige, and he wonders whether
access to digital documents will prove to
have similar symbolic value. Last, Buck-
land observes that digital documents
should lead to a national rather than a
local focus for collection building.

Information Subsidy
In thinking about library collections, Da-
vid W. Lewis argues that it is important
to consider not what libraries do, but
rather what they are for. His answer
to this question is that libraries are the
means that communities and organiza-
tions use to provide information subsidy.
Subsidy is important, he argues, because
institutions and communities recognize
that, left to their own devices, individu-
als will not and cannot acquire and use
all the information they need to be fully
productive. Institutional or community
subsidy is thus justified to create librar-
ies as a common good. Lewis points out
that, traditionally in the paper world,
libraries have done this by creating local
collections.

As information becomes digital and
moves to the network, Lewis argues that
libraries will become less important and
asks us to consider “whether libraries are
the only, or even the best, means of mak-
ing information easily and conveniently
available.” In answering this question,
he suggests that there are two forms that
information subsidies are likely to take.
The first is to purchase commercially pro-
duced content on behalf of the members
of an institution or community. This is a
familiar role for libraries even though it is
not yet seen as fully substituting for local
collection building. The second form of
subsidy will be to create free and open
content as part of the information com-
mons. This will be more complex as it
will be difficult for institutions to justify
the expense. He suggests that, with this
form of information subsidy, the library’s
role is uncertain. Lewis raises the possibil-
ity that libraries will fade as the channel
for information subsidy and asks that
libraries work to forge a future where the
information subsidy persists, even if the
organization of the library is replaced.

The Specific Drivers of Change
for Libraries and Scholarly
Communication
Taken together, Buckland and Lewis pro-
vide a useful framework for considering
the future of academic library collecting.
We are quickly approaching the time
when collections, at least as they serve
the dispensing role, will be largely digital.
Buckland explains why this is when we
can expect significant changes in library
practice to occur. Buckland also draws
our attention to the distinct preservation
and dispensing roles and by implication
points out that there are likely to be very
different strategies for each. Lewis’ focus
on subsidy reminds us of the continuing
need to make social investments to create
information as a public good. He points
to the need to look for the most efficient
ways of doing so and suggests that, in the
digital world, open access will be an im-
portant but politically complex strategy.

While, as noted above, the general
cause of the changes academic libraries
will soon have to make is the development and maturity of computational and network technologies, there are a number of specific developments relating to libraries, books, and scholarly communication that drive change and provide the building blocks upon which new library practices will be constructed.

1. Digitization of Content

The first large collections of digital content were the full-text journal aggregations created by Lexis/Nexis and InfoTrac in the 1980s. The successors to these early projects now provide access to tens of thousands of current journals and newspapers, many with significant backfiles. JSTOR was founded in 1995 in part to help relieve the space pressures faced by many academic libraries. In February 2011, the collection contained the full runs of 1,371 journals and over 40 million pages.\(^9\) At about the same time, many journal publishers moved their current and (in many cases) backfile content onto the web. Through the efforts of publishers and projects like JSTOR, nearly all journal literature is available digitally, as are the backfiles of most core titles.

In 1994, the Government Printing Office began GPO Access to distribute digital versions of federal documents. This project has grown so that now, for all practical purposes, all recently published federal government information is available in digital formats. Project Gutenberg began digitizing literary works in the public domain in 1970 and was one of the first digital library collections on the Internet. In 1999, netLibrary began selling library access to electronic books, mostly frontlist titles from academic publishers. In the early 1990s, the Library of Congress began the American Memory project with the mission of providing “free and open access through the Internet to written and spoken words, sound recordings, still and moving images, prints, maps, and sheet music that document the American experience.” The collection currently includes over 9 million items.\(^{10}\) Many states followed this example and established their own “memory” projects and many libraries have digitized portions of their special collections.

Reasonable progress in digitizing content was being made, but it was projected that it would take decades to finish the task. Then, in 2004, an 800-pound gorilla walked into the room when Google announced its Library Project. Originally the project partnered with four research libraries and intended to digitize 15 million volumes in a decade. By March 2012, Google had scanned 20 million books and estimated that there are about 130 million books in existence. They intend to digitize all of them within the next ten years.\(^{11}\) The project raised significant copyright issues, and, within a year, publishers and authors filed lawsuits that remain unsettled. The Google Books project was the impetus for two other important mass digitization projects, the Internet Archive’s Ebook and Texts Archive containing over 3.8 million public domain titles and the HathiTrust, a collaborative program of over 70 academic and research libraries that contains nearly 10.6 million volumes, much of it copies of Google and Internet Archive scans.\(^{12}\) A study by Constance Malpas found that, as of June 2010, the median rate of duplication between titles held by members of the Association of Research Libraries (ARL) and the HathiTrust was 30 percent and that, if the current growth trajectory of the HathiTrust collections holds, it will be more than 60 percent by June of 2014.\(^{13}\)

Even though the legal issues regarding the extent and nature of the use of the Google corpus is unresolved, its existence requires libraries to change the ways they think about how books will be accessed and used in the future. Many complex issues remain, yet it seems likely that, sooner or later, most of the books in existence will be scanned. It is then only a matter of time before lawyers, the courts, Congress, or (most likely) some combination of the above will find a compromise to make at least some significant portion of them available to individuals.
and libraries. In 2010, Michael Cairns estimated that, if Google were to be able to license the copyright material in the Google Books database, the annual cost for the average academic library would be $55,000. His cost estimate for the New York Public Library, with the main library and 50 branch libraries, would be $340,000 annually.\textsuperscript{14} This prospect changes everything. It seems inevitable that the provision of documents, Buckland's dispensing function of libraries, will in large part be met by web-scale enterprises, not by individual local library collections. Libraries are likely to remain the payment mechanism for some of the access. It is also easy to imagine an iTunes-like model for individual purchases.

2. Print Repositories
Lizanne Payne's study identified, as of the summer of 2007, 68 high-density print repositories in North American academic and research libraries with the capacity to hold over 70 million volumes, or approximately 7 percent of the one billion items owned by North American academic libraries.\textsuperscript{15} These facilities provide optimal environmental conditions for the long-term preservation of paper and restricted access making them more secure than open-stack libraries. Because in most cases the collections are shelved by size, weeding these collections is cost prohibitive. This turns out to be a good thing. As Paul Gherman puts it, "If our storage collections are not weeded, then we could build a 'trust relationship,' whereby libraries with storage collections would promise to retain these in perpetuity and loan materials to others, thus allowing other libraries to make withdrawal decisions based on this agreement."\textsuperscript{16} Both OCLC and ARL have begun programs to coordinate collections in print repositories. Doing so is not without its challenges. Malpas points in particular to the lack of a robust discovery and delivery service for the collective collections.\textsuperscript{17}

The potential efficiencies of a coordinated group of print repositories with a service that provides print copies of works to libraries is significant. Paul N. Courant and Matthew “Buzz” Nielsen calculate present values for perpetual storage per book. Their figure for open stacks is $141.89 and for high-density storage is $28.77. If the high-density copy replaces multiple open-stack copies, the saving would be even larger.\textsuperscript{18} Malpas concludes that, if a robust shared print service were in place today, the median space saving for an ARL library would be 45,000 assignable square feet. The cost avoidance if a shared print service for mass-digitized books were available today would be between $500,000 and $2 million per ARL library.\textsuperscript{19} Paul Genoni and Eva Varga document similar savings for Australian libraries.\textsuperscript{20}

It seems likely that legacy print collections can and will be preserved for the long term. It is also likely that large research libraries will continue to purchase hard-to-acquire and little-used items and add them to print repositories. Many academic libraries will then be able to take advantage of these collections to withdraw many books from their collections and in this way both free space and save dollars from operating budgets. There will be costs, and who pays and how much will need to be negotiated; but, however this is done, there will be significant savings over current practice.

3. E-book Readers and Print on Demand
The first e-book readers were introduced in the late 1990s. They needed to solve two problems. First, they needed to create a satisfactory reading experience; and, second, they needed to embed digital rights management into the product so that publishers would release their content into the e-book market. It took nearly a decade, but in November of 2007 Amazon released the first version of the Kindle. It sold out within hours. In late 2009, Barnes & Noble introduced the Nook. Apple released the iPad with an e-book app in April of 2010 and sold 3 million devices in the first 80 days.\textsuperscript{21}
In July of 2010, Amazon announced that e-book sales for the Kindle topped hardcover sales. Ten months later, in May 2011, Amazon sold more e-books than print books. While the printed book is far from extinct, many users find the e-book reader experience to be as good or better than a paper book, and the digital rights management systems have satisfied publishers. It is clear that e-books and e-book readers have become an important part of the mainstream market for books.

With the development of digital printing processes, it became possible to reduce print runs significantly, finally allowing printing to be done one item at a time. The ultimate expression of print-on-demand (POD) may be the Espresso Book Machine (EBM). As On Demand Books, the company that produces the EBM, puts it on their website, it is designed to be “the next generation of POD technology: a fully automatic, low-cost device that could be placed in a neighborhood bookshop, coffee shop, newsstand, library, hotel, even aboard a cruise ship or in airports.” An EBM costs a bit more than $100,000 depending on the printer with which it is paired and can print and bind books for a penny a page. On Demand Books has arrangements with a variety of providers that manage rights for books printed on the EBM.

The combination of good e-book readers and POD technology that can be deployed in local libraries and neighborhood bookstores means that the distinction between paper and digital becomes unimportant. Content will be created and delivered digitally, but as Lewis puts it, “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 created and delivered nearly as quickly, cheaply, and reliably as digital files.”

4. Open Access Publishing
As defined by Peter Suber, “Open-access (OA) literature is digital, online, free of charge, and free of most copyright and licensing restrictions. OA removes price barriers (subscriptions, licensing fees, pay-per-view fees) and permission barriers (most copyright and licensing restrictions).” There are two primary open access mechanisms. The first is Gold OA wherein articles, often peer reviewed, are published in web-based journals where costs are covered by author fees, institutional subsidies, or grants. The second mechanism is Green OA and involves the deposit of articles and other content in institutional or subject repositories. Together these two mechanisms have made a large body of scholarly content freely available. A report by OutSell, a market research firm focusing on the information industry, concluded that 9.6 percent of peer-reviewed journals and 6.1 percent of non-peer-reviewed titles were open access in April 2009. Mikael Laakso et al. conclude that 7.7 percent of scholarly articles were open access in 2009.

Open access began in the early 1990s and has grown at a faster rate than the scholarly literature generally. Laakso et al. estimate there were 20 open access journals publishing 247 articles in 1993. By 2000, the numbers were 19,521 articles and 741 journals, and, in 2009, 191,851 articles and 4,767 journals. More recently Laakso and Bjork estimate that in 2011 340,000 articles or 12% of all scholarly articles, were published in 6,713 open access journals. The Directory of Open Access Journals (DOAJ) documents similar growth. In 2004 the DOAJ listed 1,133 titles; and, by January 2013, that number had risen to 8,600, with annual increases in the number of titles of between 17 percent and 30 percent.

Lewis argues that Gold OA is a disruptive innovation as defined by business theorist Clayton Christensen and, as such, the growth of Gold OA can be expected to follow an S-curve, not a straight line. Using methods outlined by Christensen for calculating the expected trajectory of the S-curve, Lewis estimates that, “Gold OA will account for 50% of the scholarly journal
articles sometime between 2017 and 2021, and 90% of articles as soon as 2020 and more conservatively by 2025." If Lewis is correct, then a major change in scholarly publishing will occur in the next decade. The change will mean that academic libraries will escape from the overwhelming burden of high and rapidly increasing journal prices that have plagued them for the past four decades. But, as Lewis notes, one of the impacts of Gold OA as a disruptive innovation is that it disrupts not only the competing business model (in this case, subscription journals) but the whole commercial system. Part of the disrupted commercial system will be academic libraries that will not be a necessary part of the OA scholarly journal system.

While open access as a business model has been focused on journals, a number of academic presses, including the university presses at Michigan, California, and Pittsburgh, have explored open access for books. Rice University reestablished its university press using an open access model, but it was closed within a few years. The National Academies Press, which for some time has made some of its books available at no cost on the Internet, announced in June of 2011 that they would make all of their 4,000 titles open access. This was possible because they had significantly reduced their production costs and had financial backing from their sponsors. It is not yet clear whether a sustainable business model for open access books exists, and if it does, what it looks like.

It also appears that freely available, if not always open access, reference works will replace many established reference tools. Wikipedia is the obvious example. Studies by Gail Herrera and by Jared L. Howland and his colleagues suggest that Google Scholar will be a, if not the, major discovery tool for the scholarly journal literature in the near future. Google Scholar’s cited reference function could replace Web of Science, and its wide coverage is likely to threaten at least some discipline-based indexing and abstracting tools.

5. Challenges to University Presses and Scholarly Book Production

The recent Association of American University Presses (AAUP) report on sustainable business models begins with a statement of the obvious: “From new technologies to new economic conditions to changing relations with stakeholders, the world of scholarly communication in 2011 looks very different than it did a generation ago.” The report asserts that university presses are a “keystone species” and that their independence from their home institutions is an important necessity for editorial quality. While most university presses have managed to survive on increasing small print runs, the reality for many is that they are small-scale, often isolated, operations with very limited capacity for innovation. Absent increased subsidy from their parent institutions, the ongoing survival of many university presses seems in doubt. In light of this, the conclusions of the AAUP report that university presses share experience, partner with each other, discuss open access, and get more external funding, seem inadequate.

It is interesting to contrast the AAUP report with the 2007 Ithaka study, University Publishing in the Digital Age. The Ithaka study was initiated to study university presses, but it evolved into the broader look at the function of university publishing. The report’s assessment was that it was the function of publishing, not the institution of the university press that was critical. The report argues that it is in the interest of universities to make a renewed commitment to publishing by marshaling all of the resources in their libraries, presses, and technology organizations. If they do so, they will “more fully realize the potential global impact of their academic programs, enhance the reputations of their specific institutions, maintain a strong voice in determining what constitutes important scholarship and which scholars deserve recognition, and in some cases reduce costs.”

It is not clear that many universities are prepared to make the required adjust-
ments in organizational structure and funding, and it is hard to be optimistic about university press book publishing in the near term. What is clear is that the downward pressures on the traditional academic book publishing market will continue. Libraries will purchase fewer scholarly books, and young scholars (and those in disciplines where publishing costs are high or where sales can be expected to be low) will have an increasingly difficult time getting a book published. This will impact academic careers. The only saving grace may be that, as university presses fail and the traditionally published scholarly book becomes a more and more difficult prospect, young scholars will be more inclined to take risks and explore alternative means of communicating their work.


As Gideon Burton puts it, “Scholarship today cannot be considered a responsible and responsive knowledge system by continuing to operate upon the assumptions or within the formats and procedures that have characterized print communications.” The Ithaka report attempts to describe where scholarly communication is heading. It will, the report says, be driven by several factors:

1. Everything must be electronic;
2. Scholars will rely on deeply integrated electronic research/publishing environments that will manage the full range of research activities from the beginning to the end of the process;
3. Multimedia and multiformat delivery will be important; and
4. New economic models will be developed.

It is hard to deny that these factors will be important, but they do not fully capture what is most likely to drive changes in scholarly communication in the next decade or two. The critical factors will be openness and social productivity.

While some would argue that it is not a scholarly tool, Wikipedia is a good example of how openness and social productivity can be used to create a knowledge resource of high quality and with network scale reach. Since its creation in 2001, Wikipedia has evolved processes that have allowed more than 77,000 active contributors to create 22 million articles in 285 languages. It has not been a trouble-free rise, but it has become the encyclopedia of choice for much of the world, attracting 470 million unique visitors a month as of February 2012. Encyclopedias are a special kind of scholarly product, and Wikipedia may be a special case. Nonetheless, it shows that remarkable things can be accomplished using openness and social production.

Burton provides a good structure for thinking about openness in the scholarly context. Open scholarship begins with open access, but goes on to open up other aspects of the scholarly communication's process. These include:

1. Open review, most likely post-publication. While the particular processes for doing this have not yet been fully worked out in the scholarly context, there are a variety of strategies that have been successful in other web contexts.
2. Open dialog. Discussion is the standard mode of operation on the web, but as Burton notes, “Traditional scholarly publishing pays homage to the ‘dialogue’ of ideas in that metaphorical sense, but in reality it soundly rejects interactivity.” This change may be uncomfortable, but it will lead to better scholarly conversations and a greater public impact of scholarly work.
3. Open process. This involves scholars being open about what they are working on and sharing early versions of work.
4. Open formats. To make sharing easy and cheap, standards and formats need to be open.
5. Open data. The data means reusable research results. Recent moves by funding agencies sug-
gest this aspect of openness may soon be a compliance issue.

Burton recognizes the challenges openness presents the academy. The academy wants what the web offers, but does not want it to change established practice, especially in regard to authentication, dissemination, and the way scholarly expertise is exercised. You can’t have it both ways, and, as Burton puts it, “It comes down to this: the more academia wishes to enjoy the benefits of the digital medium, the less it can hold on to restrictive and closed practices in the production, vetting, dissemination, and archiving of information.” Simply put, traditional scholarly structures and practices are elitist and based on exclusion. The processes and practices of today’s digital world assume inclusion and only work well when it is a core value. In the past several years, a number of studies have clearly documented enhanced productivity and a greater pace of discovery in situations of open versus closed scholarship. There is little doubt that openness works.

Social production grows out of openness. It is the capacity, made possible by the Internet, to organize tasks so that broadly dispersed individuals can contribute to a project. It also allows a range of contributions. Some people will make only one contribution; others will make many. In many cases the contributions are voluntary, though often organizations contribute the work of their employees. Social production works because the Internet reduces the cost of communication and coordination by orders of magnitude so that formal organizations are no longer required to get things done.

Examples include open source software projects like Linux or Apache, as well as the aforementioned Wikipedia. Other examples are Ushahidi, a system to collect, visualize, and map data that was developed in response to the post-election violence in Kenya in 2008; and eBird, a project of the Cornell Lab of Ornithology, which provides tools for birders and uses the resulting data in research. SETI@home, which uses free cycles on volunteers’ home computers to search for extraterrestrial intelligence, is a form of social production as well. In a well-publicized example, several groups of gamers solved a complex protein folding problem involving a retrovirus enzyme important to AIDS research whose solution had eluded scientists for a decade. Using Foldit, a game environment designed for this purpose, the gamers found the solution in ten days.

Yochai Benkler asserts that “social production is a real fact, not a fad. It is the critical long term shift caused by the Internet.” Exactly how openness and social production will play out in the academy is still unclear, but it will happen. At least at the beginning, many projects will be outside the established channels of distributing and preserving scholarship. Managing this content will be a challenge that academic libraries should tackle.

2020

To review, by 2020 in all likelihood the following things will be true:

1. Google will have made good progress in its project to digitize all of the world’s books. Nearly 100 million volumes will have been digitized. Most of these will be available through Google, the HathiTrust, and the Internet Archive in the United States and national libraries in other countries. Most public domain works will be freely and easily available on the Internet.

2. Google, authors, publishers, and all other interested parties will have reached a court-approved process for charging for the use of in-copyright works. Mechanisms will include modest pay-for-use fees for individuals and manageable site license arrangements for libraries. Frontlist e-books are available as they are today.

3. All published content will be delivered electronically. Most
readers will use e-readers or their tablet computing devices. When a printed book is desired, print-on-demand will be easily and cheaply available.

4. Print repositories managed by several consortia of research institutions will house multiple copies of all of the printed published works previously housed in academic libraries’ open stacks. Use copies can be loaned, while other copies are maintained as preservation copies of last resort. The national legacy print collection will be safely preserved.

5. Open access will be well on its way to becoming the dominant model for scholarly journal publishing. Because of this, the pressure on library serials budgets has been declining for the past several years. Universities are making other investments to support open access journals and monographs.

6. Many university presses have failed, and those that remain have been reorganized into broader university publishing units where the library is often the lead organization. Research university investment in the publishing function is stable and its value is recognized.

7. A variety of new open, often collaborative, means of scholarly communication exist. They are web-based and media-rich. Works evolve over time based on dialog. Mechanisms for peer review and evaluation are beginning to consolidate. Libraries will have recognized that they should be playing a role to support these ventures and in preserving the results. They are beginning to understand how to do so.

**Academic Library Collecting in the New World**

Given the very different world in which academic libraries will find themselves ten years from now, it is important to begin considering what changes in practice will be required. Scott Walter argues that academic libraries need to move beyond defining themselves by their distinctive collections and move to definitions of excellence based on distinctive services. But until we change how we do our collecting, there will be little room for other services to become distinctive. Current practices have opportunity costs; and, unless these costs are captured and reinvested to create services that are distinctive, academic libraries risk becoming expensive and irrelevant. It will be important, though difficult, to get out in front so as to avoid a downward spiral of outdated services and collections leading to reduced funding leading to limited capacity, etc.

All academic libraries will need to pull back from their reliance on local collections as the core of their strategy for providing users with documents. Different libraries will need to work out their own particulars; but, at a minimum, all will need to do the following five things.

1. **Deconstruct Legacy Print Collections**

The Taiga Forum frames it nicely: “Within five years, graduate students and faculty will fill all their information needs online, never coming into the library, yet they will continue to idealize the library as a sacred place to commune with books. Libraries will respond by flipping their stacks into designer reading rooms that use books as decor.” When all of the books, journals, and image and manuscript collections that matter are digitized and paper copies are safely stored, local collections will no longer be required for either the preservation or dispensing function. But what do academic libraries do with all of the volumes they have on their shelves now?

The first step, which should begin today, is to radically slow the growth of print collections. It makes little sense to add material now that you will regret having in a few years. Government documents and print journals are the easiest
place to begin. Books will be more difficult. Some libraries are slowing the additions to their print collections. Few are doing so as aggressively as they should.

Having limited the growth of print, the next step is to begin reducing existing print collections. Again, journals will be the easiest places to begin. In most cases, the paper copy has little artifactual value, in many cases there will be reliable digital copies in collections like JSTOR, and bibliographic record changes will be simple and inexpensive. Government documents appear to be a good second target of opportunity, though projects to digitize the complete historical collection are still in progress and coordination of print repository collections will require action at the state or regional level. Reducing book collections will be more difficult politically and logistically. Individual libraries will want to assess the speed with which they proceed. Strategies will need to be developed that take into account the costs of changing bibliographic records, circulation patterns, and the availability and cost of acquiring backup print copies. While the politics of withdrawing books might seem perilous, the fact that it will be cheap and easy to acquire print-on-demand reprint copies should serve to encourage bold action rather than hesitation. A number of engineering and medical libraries are already pursuing bookless strategies, so there will be experience to build upon. Given the cost of storing books documented above by Courant and Nielsen, the savings to libraries will be significant; but, for better or worse, many of these savings are not likely to accrue to the library. Rather, the campus will take over the freed-up space and repurpose it.

2. Move from Item-by-Item Book Selection to Purchase-on-Demand and Subscriptions

Three things have long been understood about the way library book collections are used. First, past use is by far the best predictor of future use. As Herman Fussler and Julian Simon put it in their comprehensive study over 50 years ago, “the variable of past use is sufficiently powerful that for libraries with 20-year use records the objective characteristics make little contribution.” Second, use of an item declines with time. Finally, use is distributed over titles unevenly and most of the use is concentrated in small percentage of the titles. This pattern is often framed as the 80/20 rule where 80 percent of the use is accounted for by 20 percent of the collection. While these patterns were well understood, they had little impact on library practice. Daniel Gore’s very reasonable proposal, made in the mid 1970s, for actively managed academic library collections as a means of limiting collection size was completely ignored. Despite the fact that purchasing multiple copies of heavily used books would clearly have increased users’ ability to locate wanted titles this was rarely done. Libraries nearly always bought as much of the little-used long tail as they could afford. These research collections satisfied faculty and added to institutional prestige but got little use. In a paper world, this could be justified. It was an insurance policy against the future unavailability of material that might someday be important.

Finally, in the paper world, selection mattered. Selecting items for inclusion in the library’s collection was a professional task. Approval plans made some inroads and created some efficiencies, but in general it was believed, not without justification, that librarians who knew their users made the best book selections. The digital world is different. As noted above, all content will be delivered digitally, and it can be provided to the user either as a digital file or as a printed book. The most important attribute of digital content from a collections standpoint is that you don’t have to own an item before a user wants it. Digital items can be acquired in real time. The library will be able to decide the universe of books it wishes to make available to its users, but it does not need to purchase the item until an actual need arises. Nor will multiple demands for the
same title create conflicts. The user will be able to get the item anywhere instantly. This resolves two historic impediments to the use of print collections: the time required to come to the library to get an item; and the uncertainty about whether or not an item will be available. These two impediments have conditioned the use of printed book collections, and it is reasonable to expect that, when they are removed, book use will increase.

Patron-driven acquisitions (PDAs) systems were first used with netLibrary e-book purchases in the late 1990s. Despite the fact that they were clearly more efficient, they were not widely adopted. The case for PDAs seems clear: every purchase is based on a certain need; therefore, every purchase leads to at least one use. Since past use is the best predictor of future use, books purchased based on what one user wants should, on average, be used more going forward. Lynn Sutton compared PDA selection versus that of librarians and concluded, “Professional collection development librarians have limited success in projecting the interests and demands of their users as measured by collection usage.”

PDAs are not without problems. If the universe of possible purchases is not well constructed, users often purchase items outside the scope of library collecting, or they may purchase too many items and exhaust the library’s budget. A key disadvantage to PDAs may be that they will be too efficient; as a result, publisher income, especially for low-use scholarly monographs, will be reduced, threatening an already weak sector. Academic publishers have counted on libraries purchasing copies in anticipation of use; if PDAs are widely adopted, this will no longer be the case.

Subscriptions may be an effective means of managing risk for both libraries and publishers. Libraries will limit their exposure to increases in use and publishers will have a guaranteed income stream. Libraries will indirectly benefit from the security that subscription models provide to publishers if this results in an increase in the output of scholarly content. The danger of subscriptions for books is that, as was the case with journals, libraries get locked into long-term commitments with unreasonable price increases. It will be vital for libraries to maintain options for both PDAs and subscriptions so that they can manage their costs and risks.

Both PDAs and subscriptions change the role of librarians in the selection process. With both, librarians will establish the overall structure of the plans, but it will not be involved in day-to-day operations. In addition, both PDAs and subscriptions should reduce acquisitions and cataloging costs. Libraries will want to capture these savings so they can be invested elsewhere.

It is currently unclear whether or not scholarly monographs will follow journals into an open access model. Few academic presses seem interested this approach, but it could have the same efficiencies for monographs that it does for journals. Libraries should monitor this development and, where possible, support it.

3. Manage the Transition to Open Access Journals

Open access will become the dominant mode for scholarly journal publishing within the next decade, but getting from here to there will present some challenges. The biggest of these will be moving away from subscription titles with declining use while still providing access to articles from these titles when that is required.

The first step is to abandon “big deal” subscription packages. These arrangements provide large blocks of content but lock in multiyear price increases and provide limited ability to save money by cutting little-used titles. Librarians have been ambivalent at best about the “big deal” since its inception. In 2001, Kenneth Frazier compared the big deal to Alice tumbling down the rabbit hole and urged librarians not to sign on. Frazier correctly diagnosed the difficulty that libraries would have getting out of “big
deals.” Recent experience indicates that this may be less problematic than librarians imagine. Claudio Aspesi, a business analyst, downgraded his rating of Reed Elsevier based in large part on his expectation that the time of the “big deal” was over, the result, he said, of “three trends overlapping: a long term unsustainable trend, a cyclical funding crisis and a more tough minded and analytical community of librarians.”

Once the “big deal” is undone, libraries need to constrain, or even reduce, their expenditures on subscription journals. This will require a regular cycle of journal cancellations. This is easier in times of fiscal stress, but libraries will want to pursue cancellations even in better times. Key to the strategy will be good use data that allows the establishment of valid cost-per-use figures. This, paired with quick interlibrary loan or document delivery, should make such cancellations acceptable. It also seems likely that the document delivery-per-article costs charged by for-profit publishers, often more than $50 per article, will cause sticker shock for faculty and thus lessen their opinion of these publications and publishers. The goal should be to limit expenditures for expensive subscription journals. Their days are numbered and libraries should not invest any more than is absolutely necessary in them. As more and more open access titles move into the top tier, it will become easier for faculty to change their publishing practices. There is some evidence that faculty authors are willing to publish in open access journals if they are prestigious, especially if they are working with international collaborators. It is therefore in the interest of libraries to continue to support open access initiatives: institutional deposit mandates; support for open access journals; or funding of open access author fees.

4. Curate the Unique
While libraries will be purchasing collections well into the future, this activity will become more routine based on either subscriptions to large blocks of content or to PDAs. Increasing amounts of content will be open access and will be discovered and delivered by web-scale services. Building collections of published materials will decline in significance. Lewis has suggested that, by 2025, this part of libraries activities might decline by half and be replaced by the activities to curate unique content that is produced on or is of interest to the campus. Some of this will be an extension of traditional special collections, but the majority of the new activity will be curating digital content. Tschera Harkness Connell has recently shown the value of this activity even as it applies to the work of undergraduates.

Lewis identified three challenges in changing to this focus. The first is these is to develop the human and technology infrastructure so that libraries have the required capacity. The technology appears to be a manageable problem. The human resources also seem to be developing. Digital collections and scholarly communications librarians are increasingly common positions in academic libraries. The part of the human resources transition that may prove most difficult will be to get subject librarians to drop traditional collection building activities and replace them with activities that engage with faculty to build digital collections. The active engagement of subject librarians in these programs will be essential. Kristi L. Palmer, Emily Dill, and Charlene Christie found little interest in moving beyond traditional roles in their 2006 survey; but since then there has been much conversation about changing roles, and we can expect actual change to follow.

The second and third challenges identified by Lewis are demonstrating the importance of curation of local content to faculty and the campus. This will certainly be helped by data management mandates from granting organizations, but libraries will want to continue to make the case for their engagement in this part of the research process. It will also be important to balance curation activities
across academic programs. Libraries may support Chemistry in different ways than they do Religious Studies, for example, but it will need to do something for both.

5. Develop New Mechanisms to Fund National Infrastructure

As resources and services move from individual organizations to web scale, more robust funding and shared governance structures will be required. Preservation organizations like the HathiTrust, Portico, and LOCKSS will require funding from the community. Services to provide print books from repositories will need either membership fees or item-by-item charges. Organizations that create infrastructure such as DuraSpace, the developers of DSpace and Fedora, the Public Knowledge Project, the developers of the Open Journal System, or Kuali OLE, which is creating an open source library system, all require funding from institutions that use their products. In addition, there are important advocacy groups such as SPARC. To date, these organizations and projects have relied on a combination of grant funding, voluntary membership fees from interested libraries, sponsorship by larger organizations, and (in some cases) partnerships with commercial firms. This patchwork of funding mechanisms is unevenly spread across institutions, with the large research institutions currently carrying the largest burden. This structure is unlikely to be sufficient as these projects expand, become more expensive, and have more libraries dependent on them.

The diversity of higher education in the United States makes it hard to imagine a national unified funding mechanism like that enjoyed by JISC in the United Kingdom, but something approaching this level of commitment will be required.

It seems inevitable that the level of shared commitment that funding core infrastructure and services will require will change the relationships between academic libraries. Academic libraries have a long history of sharing that has been mutually beneficial. Most libraries get more or less what they give, though the contribution is small. In the near future, shared commitments could represent 20 percent or more of some libraries’ budgets. This will be a fee for service, not sharing. As John J. Regazzi documents, there has been a significant difference in the growth in the resources available to libraries of different sizes. In general, large, doctoral, particularly private universities have had increases in budgets, while other libraries, particularly small to medium public universities and community colleges, have not. The level of required commitment and the disparity in resources will challenge existing collaborative organizations, and smaller institutions may be left out.

The danger is that too many libraries will choose to be free riders and not pay their share. In fact, it is likely that many campus leaders will see this as a rational response to the changing environment. It will be important for library directors to begin making appropriate contributions and to begin talking to their campus leadership about the importance of doing so.

It may be that a structure something like the United Way would be useful to collect contributions and make decisions on the degree of support for individual projects. Contributing to a single trusted agency that in turn makes judgments about how to allocate funding to particular projects could make investments in national projects easier to justify, especially for smaller libraries.

Conclusion

By the early 2020s, it is easy to imagine the following: Print collections will have been reduced by at least half in most academic libraries. The space will be used for a combination of enhanced reader spaces and other activities. Many libraries will have reduced the amount of their budget to build collections by purchasing published content. This savings will accrue from reductions in materials budgets and from a decline in the amount of staffing, both professional and clerical, required to select, acquire, and catalog locally held...
material. The published content that is purchased will be acquired in large blocks on a subscription basis or on an item-by-item basis in a PDA model. Lewis has suggested that the combination of a PDA and reduced print collection suggests that it would be cheaper to let the user keep the item so the library would avoid the future cost of storage.66 This approach might be extreme, but it is not unimaginable. Open access will be the dominant model for scholarly journals and will represent a growing portion of scholarship of all kinds. New vehicles for scholarship that use openness and social productivity will be evolving, and librarians will be playing a role in their development, use, and preservation. Some of these projects will be supported by web-scale organizations, others by individual libraries.

There will be two keys to success in this new scholarly communications ecology. The first will be the willingness of universities to continue to invest, in large part through their libraries, even as the mechanisms to do so change. Second, librarians must embrace new roles and abandon old practices. Collections and collecting must be transformed. If these two things happen, the scholarly communications system in 2025 will be healthy, and ultimately that is what matters.

Notes
3. Ibid., 55–56.
4. Ibid., 57.
5. Ibid., 57–58.
6. Ibid., 59.
8. Ibid., 192.
31. The lesson many have drawn from the Rice University Press failure is that open access is not a viable model for academic book publishing. I would argue that the Rice case is not sufficient to prove this proposition. The one clear lesson from the Rice example is that one should not expect an engineering school to support a humanities-focused university press.


41. Ibid.

42. Ibid.


44. Ushahidi, available online at www.ushahidi.com/ [accessed 24 January 2013].

45. eBird, available online at http://ebird.org/content/ebird [accessed 24 January 2013].

46. SETI@home, available online at http://setiathome.berkeley.edu/index.php [accessed 24 January 2013].


48. Yochai Benkler’s book The Wealth of Networks (New Haven, Conn.: Yale University Press, 2006) is one of the best explications of social production. This quote is from his July 2005 TED Talk “Open Source Economics,” available online at www.ted.com/talks/yochai_benkler_on_the_new_open_source_economics.html [accessed 24 January 2013]; the text can be found online at http://blog.ted.com/2008/04/16/yochai_benkler_1/ [accessed 24 January 2013].


53. See: Fussler and Simon, Patterns in the Use of Books; Michael H. Buckland, Book Availability


64. Historically, the acronym stood for “Joint Information Systems Committee.” JISC is funded by all the U.K. post-16 and higher education funding bodies as well as the Research Councils. This is its mission: “To provide world-class leadership in the innovative use of Information and Communications Technology to support education, research and institutional effectiveness.” This includes a variety of library related programs. See the JISC website online at www.jisc.ac.uk/aboutus.aspx [accessed 24 January 2013].


PsychiatryOnline.org offers the most comprehensive online access available for psychiatric textbooks, journals, and professional development tools. This virtual library provides psychiatrists and mental health professionals with key resources to assist with diagnosis, treatment, research, and professional development.

Request a free 30-day institutional trial at www.PsychiatryOnline.org/freetrials

Or email institutions@psych.org for more information.