SEARCHER

Volume 7, No. 3 • March 1999

Building Earth's Largest Library: Driving into the Future

by Steve Coffman

Director, *FYI, County of Los Angeles Public Library*

See Steve Coffman's **follow-up article** in the July 1999 issue of *Searcher.*

It's been called, "The click heard round the world."

It is the quiet sound of people all over the planet buying their books on Amazon-dot-com. Certainly it can't match the import of the battle of Lexington, and it lacks the drama of the storming of the Bastille, but for those in the book trade, the rise of Amazon has been every bit as revolutionary, and the click of that mouse sounds as loud as the roar of any cannon.

And it all happened so fast. Less than 4 years ago, as the story goes, Jeff Bezos and his wife were driving across country with all of their possessions in the back seat. He had just left his job as a securities analyst for a Wall Street firm to try to get into something on the Internet, but when they started out from New York, he wasn't exactly sure where they were going, nor what that business was going to be. However, by the time they'd crossed the Mississippi, he'd decided that the answers were Seattle and books. And the rest, as they say, is history.

As of this January, Amazon reported it was serving a total of 6.2 million customers from 150 different countries, 58 percent of whom were repeat buyers. In the last 3 months of 1998 alone, Amazon shipped some 7.5 million books, CDs, and videos — enough to fill a bookshelf 101 miles long. And based on fourth quarter sales, Amazon had achieved annualized revenues of over $1 billion just three-and-a-half years after its doors were opened. That's over 30 percent of the total revenues generated by Barnes and Noble last year with its more than 2,000 stores and 27,000 employees. Amazon has done it all with no store and only 614 employees.

Clearly, they must be doing something right.

In their 1997 Annual Report, Amazon outlines what it believes are the primary competitive advantages offered to customers doing business over the Net. While far too numerous to detail here, I would like to highlight a few of the more important ones that help explain why the Amazon model has become so successful.

Selection

Amazon offers a selection of over 3 million titles, including all 1.5 million English language titles currently in print, as well as everything listed in *Books Out of Print*. That's more than 17 times the 175,000 to 200,000 titles available at your local Barnes and Noble or Borders and way more than the 20,000 to 40,000 titles you might expect at a typical neighborhood bookstore. As Amazon's ads have pointed out, it would be impossible to

duplicate that kind of a selection in any physical building — except perhaps for the Pentagon. The upshot is that no matter what type of book you want, you more likely can get it from Amazon than from any "real" bookstore.

Of course, Amazon doesn't actually stock all those titles. It maintains an inventory of about 300,000 of the most popular items in its warehouses in Seattle, Delaware, and Nevada. These books list in their catalog as "Ships Within 24 Hours." Then they have arrangements with Ingram and other wholesalers (more wholesalers now that Barnes and Noble has bought Ingram) to deliver another 420,000 titles that they don't have in stock within 24 to 48 hours. The rest of the in-print material they order directly from the publishers, with most of these titles listed in the catalog as shipping in 2 to 6 weeks. They use Powells and several other large bookstores and book search services to deliver the out-of-print material; these titles list in the catalog as "hard to find."

Ironically, what Amazon has really created here is a special order service, similar to what you could get at any good bookstore. Amazon actually has less than 10 percent of the 3 million titles advertised in stock; the rest Amazon will try to get for you if you want. That's really no different than the service you could get at B&N, Borders, or any good independent bookstore. ("If we don't have the item in stock, we will gladly special order it for you.")

From old-style bookstores, this kind of special service never really got much use because you pretty much had to know exactly what you wanted before you could place an order — either that or you had to be able to pick it out of Books In Print. What Amazon did differently was to integrate listings for all of the books they could get — whether they actually owned them or not — into a single catalog that did not distinguish between what they held and what they would have to order. It only noted differences in the time it took to ship the material — which, of course, is all that really mattered to the customer anyway. And with that, Amazon transformed a traditional, low-use, special-order service into Earth's Largest Bookstore.

Accessibility and Convenience

Real bookstores are located in real places and stay open real hours. And if you happen to live too far away from them, or you'd like to use them when they are not open, you are just out of luck. Not so with Amazon, of course. Amazon opens shop everywhere an Internet connection exists. Increasingly that means right on your desk at work and in your study at home and anywhere in the world you live. Best of all, it's convenient. It waits patiently for you to come use it whenever you want, whether it's 2:00 p.m. in New York, 11:30 at night in Bangkok, or 4:00 a.m. Sunday morning in Mexicali. You don't have to do anything special to use it. You don't have to get in your car, or find a parking space, or put up with surly, underpaid clerks. Just fire up your computer and click. You don't even have to get dressed if you don't want to — after all, nobody on the Internet knows you're naked.

Of course, this kind of arrangement also proves very convenient for Amazon, because they don't have to fork over all the millions of dollars it costs to build and maintain hundreds of brick-and-mortar, physical stores, nor do they have to pay the salaries of all the clerks and managers it would take to run them — especially not 24x7 (24 hours a day, 7 days a week). That's money they can keep for themselves or maybe give back to you in the form of cheaper book prices to make an even stronger case for doing business on the Web.

Quality of Web Site and Search Tools

Amazon has figured out how to recreate on its Web site many of the features people like best about traditional bookstores. Books are organized into broad subject categories that you can browse through on the Web, much as you might browse the shelves at a Borders or Barnes and Noble. In a physical bookstore, however, books can only be arranged one way. But on the Amazon site, you can find a range of organizations for the book

listings, all reached with the click of a mouse. Don't want to browse the titles by subject? Click here and browse through lists of books that have won awards, or books that might make good gifts, or books to fit your mood on a dark and stormy night. That's one real advantage that electronic catalog records have over real books, the ease of rearranging listings to suit different customer needs.

The classified arrangement of traditional bookstores can sometimes get in the way when in trying to find a particular title, customers can't imagine where the bookstore managers hid the book. Not so at Amazon. If you prefer to search for specific books, rather than browse the classified sections, Amazon has a wonderfully fault-tolerant catalog that allows you to search by author, title, subject, and all the other standard ways, plus some ways not so standard — children's literature by grade level, for example. Not only that, if you make a mistake and misspell the name of the author or get the title wrong, Amazon will try its best to figure out what you meant and give you the most likely matches.

But a good catalog, no matter how nicely arranged, is no substitute for actually hefting the book, checking out the cover art, scanning the table of contents and the jacket blurbs, and maybe even cracking the cover, smelling the pages, and reading a chapter or two — something you can do in any bookstore. Amazon hasn't figured out how to handle the new book smell yet (before you buy it anyway), but it has managed to duplicate much of experience of examining a book by fleshing out those bare-bones catalog records with all kinds of information and content. Many of the titles on the Amazon site have at least a picture of the book jacket, some cover blurbs, full text of any reviews from the primary reviewing media, customer reviews, author interviews, table of contents, and some selections from the text. While that may not exactly make up for actually holding the book in your hand, it can come pretty close, and in some cases — full text of reviews, for example — Amazon actually provides more information on the title than you could find in the typical bookstore.

Personalized Service

Personalized service. Sounds like an oxymoron on a Web site where you can't even talk with any live human being, to say nothing of a knowledgeable bookseller. But the fact of the matter is, Amazon has managed to take advantage of technology to offer personalized customer service that equals and sometimes even exceeds the care you would get at many of the best independent bookstores. The only difference is that Amazon doesn't have to employ thousands of bookish clerks to provide it; Amazon's computers do the job instead. They remember who you are and welcome you to the site when you logon. (How many booksellers do you know that know you by name?) They remember what you have bought and will send you nice little e-mails notifying you when new books have been published in your subject areas or by your favorite authors, or when the next book in the series you are reading has become available.

They have another system that will recommend books for you based on what others with similar preferences are buying. For example, if you purchase Tom Sawyer by Mark Twain, Amazon also recommends the Adventures of Huckleberry Finn, The Prince and the Pauper, and Life on the Mississippi, as one might expect for a Mark Twain aficionado. But they also recommend Dracula, Around the World in Eighty Days, Gulliver's Travels, and — for the more musical among you — a CD called Zoot Suit Riot, the Swingin' Hits of the Cherry Poppin' Daddies. Even the most attentive book store clerk might have missed those last four, but the computer just tracks the most frequent purchases of people who liked Tom Sawyer.

Of course, Amazon will also help you pick out gifts, wrap them, enclose a personal note, and ship them anywhere in the world, just like any good bookstore. Finally, if you don't want any of this extra service, if you would rather just remain anonymous, all you have to do is turn it off. Just try doing that with a bookseller once he thinks he's got your number.

Add this all together and toss in few other factors we haven't discussed here, like discount pricing, great customer service, and fast delivery right to your home or office, and those 5.5 million customers and over a billion dollars in annual sales don't seem so surprising. The bottom line is that for a growing number of us, Amazon simply represents an easier, better, and more convenient way of buying books than anything the competition can offer — whether that competition is a 50,000-square-foot chain store with 175,000 titles, a cafe, and comfortable chairs, or a charming little independent with limited selection but knowledgeable booksellers. Amazon incorporates the best of both these worlds and more. That's why they have done so well. That's why so many of us turn to them first for buying books. And that's what makes the traditional booksellers — both chains and independents — so worried.

Perhaps, it's a good thing then that Amazon doesn't lend books, as well as sell them, or we librarians might be in really big trouble.

Nor is this just idle speculation. If Amazon's skyrocketing sales figures show anything, they show clearly that a large section of the book-buying market prefers the selection and ease and convenience of purchasing books from Amazon.com instead of from traditional "brick and mortar" bookstores. Studies show that many of these same people are also heavy library users. If they like the Amazon model for buying books, there is every reason to suspect they would like it for borrowing books as well.

So let's do it. Let's pretend for a minute that we are Jeff Bezos and his wife taking that long trip across the country. We've just crossed the Mississippi, and we've finally decided our future . We're still going to Seattle — always liked the weather there — only this time instead of building Earth's Biggest Bookstore, we've decided to take a little more altruistic bent and see if we can't create Earth's Largest Library. The way I drive, we've only got a few hours left, so there's no time to lose.

The Collection

First, let's tackle the collection. This one is a real natural, because libraries suffer from many of the same liabilities that Bezos saw in the traditional bookstore. Like bookstores, most libraries offer a pretty limited selection of books. Local collections are usually quite modest — a single Barnes and Noble has more titles than 85 percent of all the public libraries in the United States — and often what the libraries do own they have out on loan. So, it should come as no surprise that patrons cannot find what they want at the public library 50-65 percent of the time — a remarkably high failure rate, even for a public institution. (Sharon Baker, The Responsive Public Library Collection, *Libraries Unlimited,* 1993).

Like bookstores, libraries also offer a special-order service. We call it ILL or interlibrary loan. Theoretically, our patrons can use it to reach over 40 million items held by other libraries — far more than the few million titles that a bookstore or even Amazon.com could offer. I say theoretically, because our special order system suffers from the same problems Bezos found in bookstores — nobody uses it. Statistics show that less than 3 percent of the average public library's circulation comes from interlibrary loan (Baker, 1993). The problem is the same as in bookstores. In order to use the ILL system effectively, you've got to know exactly what you want. Since most library catalogs don't go beyond the local collection, most patrons never even find out about those other 40 million titles out there, to say nothing of actually ordering them.

The solution is the same one that Bezos came up with for Amazon: 1) Do away with the local catalog; 2) Build a global catalog that allows patrons to search all the material in the local collection PLUS the 40+ million items available through ILL in the OCLC Worldcat database, and, because the OCLC database doesn't pick up everything in print, toss in the 3 million items from the Amazon database for good measure. (We'll talk about how you might deliver those non-OCLC titles in a bit). Like Amazon, such a catalog would not distinguish

between local holdings and those that might come in from elsewhere. The patron would see a single catalog, the difference between the items displayed in predictions of the amount of time it would take to get them. Some items would carry the "available immediately" mark if they happened to sit on the shelf in the local collection; others might carry a 24-48 hours mark, or 3-4 days, or 2-6 weeks, depending on the arrangements worked out with supplying libraries.

Now there's a collection that would put even Amazon to shame. And it would make every library that used it, no matter how humble or how small, the equal of the richest and best-known research institutions in the world, just as Amazon has turned every Web terminal into a massive bookstore.

The Catalog

But just giving your patrons access to a bunch of catalog records — even 43 million of them — will not suffice to create Earth's Largest Library. If it could, Amazon would simply have slapped a copy of *Books In Print* on their Web site and called it a bookstore.

The fact of the matter is that your basic catalog record — whether BIP or AACR2 — is really an historical relic of a time when we had to fit everything we knew about a book on a 3x5 card. While this sort of bibliographic record might still work as a finding aid for books in physical collections where you can walk over to the shelf and examine them, our plain vanilla catalogs have proven woefully inadequate in helping people find and select books residing on shelves thousands of miles away — or even on the second floor.

In order to make such a massive catalog useful to our patrons, we would have to "enhance" it, much as Amazon has enhanced the basic BIP records they started with. Wherever possible we would want to include cover art, jacket blurbs, selections from the text, links to reviews, customer comments, author interviews and articles, and any other content that would help a person decide whether to request a particular book. We might even go Amazon one better and come up with lists of "library recommended" books in each category. After all, we try to identify the best quality books when we do collection development now. Readers should appreciate that kind of evaluation in a massive catalog even more.

In addition to fleshing out the content of the catalog records, we would want to arrange the records into all kinds of different browsing categories, and of course, include the various software "recommendation engines," forward e-mail suggestions to patrons, and follow all the other strategies Amazon has pioneered for personalizing the catalog to fit patron needs.

Sounds like a massive job to transform the OCLC catalog into something that looks like Amazon.com, but it might not prove as difficult as it seems. You'll note that Amazon hasn't really "enhanced" all its 3 million records. They've actually only got detailed content on a relatively modest number of their more current and popular titles. The rest of their records come straight from BIP. So, a little can obviously go a long way. Plus, many publishers have begun to realize the importance of content to online bookselling and now routinely provide cover art, tables of contents, and selection guidance for all the titles in their catalogs. Several major book wholesalers have developed Web-ready catalogs, complete with links to reviews, author interviews and articles, and all the other types of content we have come to expect from Amazon.

The point here is that "enhancing" a catalog would be much easier today than a few years back when Amazon first got started. I bet that if we could add content to a small percentage of those 43 million records and start off with a site that looked anything like Amazon.com, our patrons would be so overjoyed, they wouldn't even notice we hadn't done the whole thing. And, of course, we would continue to add content and build up the catalog as we went along. Nor would we have to use the identical enhancements that Amazon has used to promote current in-print books. We could tap into old book reviews and recommendation lists and librarian expertise for

out-of-print material, the bulk of most library collections.

Circulation System

Now that we've got all these books listed properly, we've got to have a way to check them out. Here again Amazon shows us the better way.

For many years now libraries have forked over millions and millions of dollars to a few commercial automation vendors in exchange for the most rudimentary online catalogs and circulation systems. Most of these proprietary systems run on proprietary networks and require specialized hardware and software, which often only the vendor sells and at exorbitant prices.

These systems have offered only the most limited functionality. Many still use ugly, character-based interfaces that even the librarians find difficult to use, much less the patrons. These online catalogs usually offer only the most basic search options: author, title, and subject. It was considered quite a coup when many finally started adding keyword searching a few years back. Some have only recently developed Web-based interfaces for their systems. And you can just forget about such niceties as automatic spelling correction, e- mail recommendations, "if you like this, you might like these" suggestions, and all the other customer-friendly features you find on Amazon. Library systems simply don't support them. If the rate of innovation among these vendors remains anything like it has in the past, it will be a long time before they do.

Finally, although many of these automation systems were sold on their ability to cut labor costs, circulation in libraries remains a very labor-intensive process. Patrons generally cannot simply check out books by themselves; it requires lots of trained circulation staff to handle these cumbersome systems. Even libraries that have begun to install self-checkout machines find that they must keep staff standing by to handle exceptions. And nobody has even begun to think about patron self-checkout over the Web.

Contrast all of this with the Amazon system. In the first place, you don't need to go down to the library to use it. It's available all over the world on any computer that can connect to the Web, any time of the day or night. It's entirely self-service. No reference assistance. No specialized training classes in how to search the Amazon catalog. (In contrast, "Using the Online Catalog" has been a staple of library training programs for years now.) No live customer service rep, even if you want one. Yet, judging from Amazon's sales figures, millions of people seem able to find what they want easily enough and to go through a detailed ordering process requiring addresses and even credit card numbers.

The Amazon system does much more than check out a book. It automatically takes up your order, verifies your credit card, and arranges for shipping. It locates the books you want and generates a pull order that tells the staff in the warehouses what to send or forwards your order to the appropriate wholesaler or publisher for fulfillment. While you wait for your books to arrive, it automatically generates e-mail updates informing you about the status of your order. Finally, the system records every element of that transaction, including who you are, where you come from, and the books you purchased — so it can get to know you and your habits better in order to provide you with even better and more personalized service the next time you visit the store.

About the only thing the catalog doesn't do is check the books back in, but in comparison with all the other sophisticated functions the Amazon system handles, adding that function would not seem to require much of an effort. Best of all, Amazon has built the most critical pieces of the system themselves using standard Web tools and applications. So they own it and can modify it and improve it at will without having to await the whim of some commercial automation vendor, as libraries often must.

Last year Amazon handled an estimated 23 million transactions, for 5.5 million customers in 150 countries around the world. That's more than one-and-a-half times the total circulation of Queens Borough Public — the library with highest circulation in the U.S. And Amazon did it all over the world, 24 hours per day. I'm not suggesting here that we librarians abandon our buildings and do business exclusively over the Web like Amazon. Many people still want to visit us in person; just as many wish they could check a book out on the Web. But what the Amazon model does show is that the Web and adjacent technologies have begun to offer better and more efficient ways of doing our work for us both on the Web itself and inside our buildings. If such a system can work so well for Amazon and its customers, just imagine what it mean could for libraries and for our patrons.

Costs

All right, we've put together the biggest collection in the world. We've designed a wonderful catalog, and we've finally got a circulation system that makes sense. Just a few hundred more miles to go before we reach Seattle. There's only one small problem left. How much is all of this going to cost? And how are we going to pay for it?

The short answer is — lots of money. Let's just take a look at the catalog, circulation system, and related infrastructure to begin with. According to Amazon.com's most recent Annual Report, they spent over $12.485 million dollars on "Product Development" in the 1997 fiscal year. This category includes "payroll and related expenses for development, editorial, systems, and telecommunications operations personnel and consultants, systems and telecommunications infrastructure and costs of acquired content" — in other words, all the costs associated with the development and operation of the Amazon Web site, catalog, and transaction processing systems.

Twelve-and-a-half million dollars is a lot of money. And that doesn't even count the money we'd have to give OCLC to use those extra 40 million records or some of the other enhancements we've discussed. Even without those additional expenses, $12.5 million is more than the entire budget of many libraries. Even the largest libraries could have a tough time shouldering the full costs of such a system alone.

The secret is, of course, that no single library would have to bear the full costs of developing and operating the system. Once you develop a single catalog with 43 million titles, one which includes records for most of the material in most of the libraries of the United States and one accessible to everybody over the Web, it wouldn't make sense for every library to continue to maintain (and pay for) its own separate catalog and circulation system. Rather, there should be a single catalog with a single bibliographic record for each title. Libraries participating in the system would attach their specific holdings information to those bibliographic records, much as they now do with OCLC. The only difference would be that the holding information would have to be entered at the copy level, so the library (or rather the patrons) could check things in and out, and the system would generate data on exactly how many copies of an item remained on hand and their location.

We would only need one circulation system for the entire catalog, although libraries might want to maintain their customer databases and transaction logs separately to help insure patron privacy. But privacy concerns are a software design issue, not a reason to have thousands of separate circulation systems.

All the participating libraries would share the cost of developing and maintaining the catalog and circulation system probably on some sort of annual subscription based upon level of usage. And it would likely still cost far less than libraries currently spend running their own feeble systems. Additionally libraries could count on substantial savings in staff costs, because if those 5.5 million Amazon customers can handle their own transactions, why would we need all those clerks standing behind our circulation desks? Of course, you would still need staff to re-shelve books, help people find books, answer questions, and other routine operations. But

a truly automated circulation system with patron check-out and check-in should allow you to significantly reduce your paraprofessional staff, just as ATMs have allowed banks to cut way back on tellers.

Although it is difficult to put an exact dollar figure on these potential cost savings, it is likely to be substantial, sufficient to convince enough libraries to join the system to raise the $12 to perhaps $20 million it would cost to develop and operate it. In fact if you could realize even a portion of the service improvements and cost savings the Amazon model promises, you could expect to capture a good portion of all the millions and millions of dollars libraries currently pour into the automation marketplace — with very little to show for it.

But a library based on the Amazon model introduces one new area of potential expense where the costs are much more difficult to predict and the answers much less obvious. Who's going to pay for all those 43 million books? Or, at least, who's going to pay to deliver all those 43 million books to patrons?

Theoretically, we should already have this problem solved. After all, we are discussing the same old ILL system we've had for years. We are simply proposing to put that system out there where patrons could really use it. So you would expect that fundamental issues, such as how much it should cost to lend or borrow an item, and who pays how much for what, should have been addressed years ago, right?

Wrong. Despite the long tenure of the ILL system, libraries have never really sat down and figured out what lending a book really costs, and more importantly what it really should cost, if done efficiently, who should pay for it, and how. The payment schemes we have are haphazard and inconsistent. Charges have little or nothing to do with the actual costs of providing the service. Some lending libraries charge; others do not. Some borrowing libraries pass those charges through to patrons; others do not. And when libraries do charge, costs can range all over the map from less than $1.00 to $25-$30 or more. On top of that, many libraries have devised all sorts of restrictive policies to specify which materials they will send to whom and under what circumstances — for example, many libraries will not loan material still in print, some academic libraries won't fill requests from public libraries until they've exhausted all other avenues, and so on, ad infinitum.

Up until now, we've managed to avoid having to come to terms with these issues simply by discouraging our patrons from using the ILL system. After all, as long as ILL remained 3 percent or less of total circulation, who really cared whether it operated on a sound financial footing?

Reality Check

All that would have to change if we decided to build Earth's Largest Library. The whole concept predicates on giving patrons transparent and unfettered access to the collections of all participating libraries. For that to work, the system must make sense financially. Lending libraries must be fairly compensated for their costs in providing the material. In fact, they should get some profit from the transactions to encourage libraries to develop collections others would want to use. Conversely, borrowing libraries must be able to pay those costs (perhaps with some help from their patrons).

Finally, the real costs of ILL must come down. I've seen relatively recent studies that peg the average cost of a complete ILL transaction at anywhere from $25-$30 (Measuring the Performance of Interlibrary Loan and Document Delivery Services, Association of Research Libraries, 1997). That's more than the retail price of many books, and those kinds of charges would not work for a system that depended heavily on ILL. What's more, they are totally out of line with expectations for real costs in a truly efficient system. After all, if companies like Amazon.com, Reel.com, and others can afford to individually pack and ship books, videos, and all kinds of other items out to customers for a small fraction of their tiny retail margins, there is no way it should have to cost libraries anywhere near $30 to process an ILL transaction, even allowing for the costs of returning the

item.

But no matter how efficient you make the system, it will still cost somebody something to borrow a book — although that figure should drop to much less than at present. If we follow the traditional model of free library services, then those costs should come out of the collection budget of the borrowing library. The question is, how could libraries — many of whom can barely afford to maintain their own collections now — afford to give their patrons access to another 43 million books?

Here again, the Amazon model may help provide an answer. As we now know, Amazon claims to have over 3 million titles, but it actually stocks only about 300,000 titles in its own warehouses, relying on wholesalers, publishers, and out-of-print book services to deliver the rest. In other words, Amazon operates a "just-in-time" inventory system. The 300,000 most popular items are stocked, and then other sources are relied upon to deliver all the other items — if and when customers want them.

In contrast, libraries have traditionally operated on a "just-in-case" inventory system. We stock a limited selection of popular titles, but we spend the bulk of our collection budgets trying to maintain a broad collection in all subject areas, "just in case" somebody might come in and want them. As a result most library collections behave according to the well-known "80/20" rule, where a small number of very popular items account for a large percentage of the circulation, while the vast majority of the books in the collection get little or no use of any kind.

This approach to collection development hurts our patrons in two ways. First, we waste their money buying books they don't use, just in case they might want them. And, secondly, we spend so much money buying books they don't want, we do not have enough left over to purchase an adequate amount of the material they do want. That's why those statistics show that the average person fails to find what they came to the library looking for anywhere from 50-65 percent of the time.

That is an awful failure rate. And while we once may have justified such wasteful collection management practices back when identifying and obtaining books outside local collections was difficult and time-consuming for patrons and librarians alike, we can hardly justify continuing such a "just-in-case" approach now when it has become so easy to acquire books virtually anywhere in the world and have them delivered to almost anywhere else in the world quickly and efficiently.

So, if we want to have the resources to provide our patrons access to 43 million books, than we need to shake off the old ways of doing things and manage our collections a little more like Amazon.com.

If we had a system that could quickly deliver almost any book on demand, it should free us up to focus our local collections more closely on those popular titles that account for so much of our total circulation. Now, I'm not suggesting here that we fill our shelves exclusively with John Grisham, Danielle Steele, and dozens of copies of Men Are from Mars, Women Are from Venus. We need to maintain balanced collections of key current titles in most major subject areas to accommodate patrons who need something right away (just as Amazon keeps a few core titles on most subjects in its inventory for immediate delivery). However, we should also prepare to adjust our collections to better suit patron demand — buying more copies of popular titles, e.g. — so our patrons would at least have a fighting chance of finding what they wanted when they came to the library, and cutting back on material we buy "just in case" somebody might someday come along and want it. With the money saved, we could pay on that 80 percent of our collection that doesn't circulate (much anyway), and we could help pay for the cost of borrowing less popular titles from other libraries — but only if and when our patrons asked for them.

How would you figure out what titles should go in the local collection, and which you should order on demand? Well, in many cases you would simply know that a particular author or subject matter will prove popular. But in other cases, you could leave that decision up to the patrons themselves, just as Amazon does. Of course, Amazon will automatically add known popular titles to its inventory. But in most cases the customers themselves influence Amazon's purchase decisions based on the amount of interest they show in a particular title. When a new and unknown book enters the catalog, Amazon first arranges to get it from a wholesaler or the publisher and then watches to see what happens to it. If the title begins to garner a significant number of orders, multiple copies are purchased for Amazon's own inventory.

There's no reason why the same practice wouldn't work in libraries. Call it "patron-centered collection development." Once you have given patrons access to all those 43 million items, including everything in BIP and Forthcoming Books, you could watch what your patrons do with the titles. For existing titles available through ILL, the first time a patron orders that book, you borrow it from another library. But as soon as a second or third request for that item comes in, you would begin compiling evidence for the title's popularity and make a purchase decision for your local collection.

New books would require different handling, since ILL could not offer them until some time after publication. So instead of allowing the patron to order the book directly, as they could with a title available through ILL, you could stick a little button on the catalog records for new books that would read "Suggest this Book for the Library," or something similar. If just one person suggested a title, you might not do anything about it (except perhaps to send them a note that you would try to get the book when it became available through ILL), but the minute a second or a third person suggested it, you would have a fairly good idea that its popularity might justify getting it for your own collection.

Of course, one small problem remains with this whole concept of re-focusing library collections on popular items that account for most of the circulation: If every library followed the popular path, we all might end up with much of the same popular material. Then nobody would have the other 80 percent of the stuff that doesn't move, the titles that we had promised in our 43-million-item catalog. However, for a number of reasons, I would suggest that this will not likely be a problem.

First, what's popular at one library will not necessarily be popular at another. The Chicago Public Library will likely have more titles on Chicago history than its counterpart in Los Angeles; academic and special libraries will collect different material than a public library, and so forth. Collectively then, a group of libraries, such as we envision here, will probably end up with a relatively diverse collection even when each of them focuses on buying more popular material for their patrons.

Second, with a global catalog and rapid delivery such as we propose, libraries could easily collaborate in collection development. One library might agree to develop a good collection in dogs, if another collects heavily in cats, and so on. Members of the consortium could offer each other better delivery times and lower costs than generally available, and the catalog could link all of these separate collections together into a whole. Patrons would see no difference in the global catalog we've described. Consortium requests would simply be routed to other members for fulfillment before they went to outside libraries. With this kind of arrangement, a library could justify getting a more esoteric book on cats which might not circulate enough among its own patrons, but could receive enough circulation from the consortium as a whole to justify its purchase.

Third, we've assumed in this discussion that the primary purpose of the library is to deliver books and information to its patrons. For most libraries that is true enough. However, a few large academic and research libraries have traditionally served an archival function — to gather, store, and preserve the most important

records of our civilization. These institutions currently hold a significant portion of the 43 million items that would appear in the hypothetical catalog. Such major resource collections would continue to play that role in the future. However, we could not expect them to fulfill that role and supply a fair portion of our patrons' requests for rare and esoteric material solely out of the goodness of their hearts. They will need compensation, otherwise we could not count on their contributions. That fact increases the importance of putting ILL on a sound financial footing. If revenues from borrowing libraries alone do not suffice to encourage such "warehouse" libraries, these days many of the old books in library collections have begun turning into major assets, now that Amazon and other online booksellers have re-awakened interest in the backlist. Many titles have slipped into the public domain. Large libraries could make a bundle reviving and reprinting forgotten books from their own collections. Publishers have begun to adopt print-on-demand systems that may keep books in print longer, but libraries may have the only extant copies of some of the publishers' books. Another potential income source.

Finally, if we really do manage to instill some profit into ILL transactions, we might expect to see something of a market develop around book lending. One advantage to having a single catalog and a single circulation system is that it would allow entrepreneurial libraries (or even commercial book-leasing services, such as McNaughton) to easily gauge the supply and demand for a particular title across the entire system. If the demand seemed to warrant it, a library might purchase the title and inject it into the system just to make a little money. So, a purchase that local library usage might not justify might prove sensible when seen as part of a commercial venture designed to profit from circulating items to a whole system.

In fact, when it comes to a commercial lending service, Amazon.com's nemesis — Barnesandnoble.com — might have an advantage. For once, Barnes and Noble's brick-and-mortar bookstores could offer an advantage over Amazon. Customers of a subscription or commercial lending service could pick up and drop back borrowed books at the Barnes and Noble stores to reduce the costs of shipping. And while they were there, of course, they might want to browse the real shelves. Naturally, the same strategy should work for libraries looking to encourage use of their physical facilities too.

So there seems plenty of reason to believe that the Amazon model would allow individual libraries to focus more narrowly on patron demand in their local collections, while insuring that collectively, libraries continued to acquire and provide access to the broadest possible range of material in all subject areas, from the popular to the obscure.

New Revenues

So far, we've looked at ways the Amazon model might help us cut the costs of automation, staffing, collection development, and other library operations so we could free up money to provide our patrons access to much broader collections. But there is another way to get money, and that's to make it. (Stealing it falls beyond the scope of this article.)

While many public sector institutions such as universities, museums, public broadcasting, and others have become quite adept at leveraging their assets to help generate operating revenues, libraries have lagged sadly behind when it comes to making money. However, a library with 43 million titles could open up some promising new business opportunities. First, with 43 million items, the library catalog would constitute the largest single source for books and book information on the Web — and as a result, would probably also become the first source people would check when they went looking for a book — a role Amazon.com's catalog has begun to assume today. However, once users found a book in our catalogs, all they could do is borrow it from us. If they would rather purchase it, they are out of luck. So, why not give them a choice? Why not link the catalog with one of more of major online bookstores (Amazon comes to mind) in exchange for a percentage of every sale

that we send their way. Amazon and Barnes and Noble already have these kinds of programs established, and some libraries have even linked their catalogs to them. To date, these library-bookstore arrangements have not generated much in the way of revenues — probably because our library catalogs are hardly inviting places for patrons to spend their time. But, with an "enhanced catalog," Web checkout, and 43 million titles, that perception could well change, and we could see some significant revenues from pass-through sales.

Used books could provide another potential revenue area. If libraries beefed up their collections with popular titles, they would also have to rigorously weed books whose day had come and gone. Right now, most of us either ignore the problem and let the deadwood stack up on our shelves, or we weed and throw them in the dump (not a good idea, just ask San Francisco Public Library), or we put them in some puny little book sale where we try to unload them for 25 cents each, or worse. None of these approaches makes much sense, and they don't do much for our bottom line either.

A better solution? Use that 43-million-item catalog — and all its users and the histories of their usage — to create a huge secondary market for used books. Just as video stores get rid of "previously viewed titles" by selling them off at reasonable prices, as each library weeded a title, it could mark the books it no longer needed and put them up for sale in the global library catalog where they would become available to a worldwide market. Titles that had not been sold within a reasonable period of time at standard "retail" used book prices could be wholesaled to used book dealers by competitive bid. If libraries truly began culling their collections as they should, a large amount of old stock would be sold off on a continuous basis, and used book sales could turn into a significant source of revenue.

At least one other major source of potential revenue remains — our patrons themselves. I hope that, with some of the cost savings we've discussed and with some of the new revenue sources covered here, most libraries could afford to give their patrons free and convenient access to almost any book they wanted as part of basic library services. Whether that becomes truly possible or not remains to be seen. But if, for whatever reason, libraries cannot fully cover unlimited free access for its patrons, then it should still allow patron access to the complete collection, subsidize what it can, and ask the patrons to pick up the rest of the tab where it can't. In this way, patrons themselves help to underwrite their own use of the collection and become a new source of revenue for the library.

Once all those patrons can search the catalog and check out books from home, many will no longer want to trundle down to the library to pick up their books. This opens the door to all kinds of new convenience services the library could offer patrons for an additional charge. Home delivery is the obvious example, of course, but there are lots of other possibilities. How about a book-of-the-month club from the library that lets patrons set up profiles of subject, author, or general interests that draw on the entire 43-million-item catalog? With 43 million items, no topic would be too obscure. The library could then ship out a new selection in the patron's favorite area every month — for a small additional fee, of course. You could prepare special "travel packs" with books and other background reading on different geographic locations for people planning trips. You could even ship them some light reading for their luggage — all, of course, for a little additional fee. Reluctant to invest library funds in lightweight paperbacks suitable for travelers trying to hold down the weight they have to lug? How about getting Amazon.com to offer your clients a special discount in return for guiding potential book buyers to them? Or maybe you could take out credits for hardback books you want Amazon to supply for your regular collection? Marketing history has shown that the public's desire for convenience known no bounds. As the Baby Boom generation enters its 50s and 60s, the demographics for delivered services will boom along with the boomers.

Current viewpoints see the library as something funded by taxes or a parent institution and designed to provide services to a specific constituency of patrons, e.g., the students and faculty at a university, the citizens of a city,

etc. There is another historical model for funding library services, however — the old subscription library — where patrons paid individually for library services, either on a rental-by-rental basis, much like the video store, or by annual subscription fees. I don't suggest this alternative as a principal method for library funding. Libraries serve a collective good for their communities and institutions. Their primary sources of funding should continue to tap collective sources like taxes or institutional budgets. However, if new or revived funding models can allow libraries to serve people outside their immediate constituencies, they should. If libraries serving some individuals have chosen not to supply the kinds of services that those individuals want or need, e.g., rural areas or other communities with little or no library services, then together these individuals could represent a significant market and a major potential source of revenue for any library that could find a way to serve them in a financially viable way. In fact by finding and developing such a new source of revenue from "library-generated patron constituencies," libraries could create revenues that contribute to services for their traditional direct constituencies.

Let's consider again expanded service to those direct patron constituencies. Suppose after you've calculated all the possible cost savings you could achieve, and after your spreadsheet projections have squeezed all potential new revenue sources dry, you still decide that it will cost more to deliver some of those 43 million titles than the library can afford to pay. Then you've got a choice: you can choose to charge your patrons the difference between what you can afford to pay and the full cost, or, you can choose not to provide the item to the patron, not to offer the service. When faced with this seemingly simple decision, most libraries have implicitly chosen the second scenario and opted not to provide any service that cannot be offered to patrons totally free of charge. It's difficult to see how such a decision benefits patrons, however, who then must either do without the item or go pay a much higher price to get it from some other source. If the library tried to cover as much of the cost of delivery as possible, the patron would benefit from the subsidy even if they still had to pay the difference. Even if they couldn't get it free, at least they wouldn't have to pay full retail prices.

The possibilities — and the revenue potential — are limited only by the imagination.

But for all our speculation here, a lot of questions remain unanswered as to how such a system might work and its viability. For example, current collection use studies focus on existing libraries with comparatively limited collections. What happens when we allow patrons access to a collection of 43 million items? Does the 80/20 rule still apply? Would you really need a collection of 8.6 million volumes (20 percent of 43 million) to supply 80 percent of the potential demand, or would a far smaller percentage do? Would the money saved from not stocking shelves with books that didn't circulate cover the costs of borrowing lesser-used items from other collections? If not, what other sources of revenue might help make up the difference? What would be the consequences of having the patrons themselves pick up part of the tab?

No one knows the answers to these questions yet, just as nobody really knew whether people would really take to buying books online before Bezos decided to find out. However, like Bezos, each of us well knows the inadequacies and shortcomings of our present methods of doing business. The Web seems to offer us a better way. But the only way we can answer those tough questions, the only way we can know for sure how it could work, is to try it and find out.

The Home Stretch
We topped Snoqualmie Pass a few miles back, and we're driving down the hill towards Seattle. We're pretty bleary-eyed, but think we can just make out some of the skyline ahead looming out of the morning mist. It's been a hard couple of days, but all this planning and cogitating has helped keep us awake across those seemingly endless miles of open road. And now that we know where we're going, maybe we can finally get a little sleep.

We are going to build the single largest library ever created in the history of the Earth. It will contain over 43 million books, videos, maps, manuscripts, microfilms, and other formats covering every known subject and area of interest that has attracted the attention of the human race since the beginning of recorded knowledge. More than any library ever has, it will mix the practical with the esoteric, the sublime with the ridiculous, and the sacred with the profane. It will be one grand, chaotic democracy of ideas. A kind of primordial soup of thought from which new ideas can take life.

Volumes which now lie scattered across the face of the globe in thousands of separate buildings and collections will come together in a single unified catalog accessible to all — and not one of those bare-bones catalogs you so often find in yesterday's libraries either, but a catalog designed for the selection decision, with records that carry reviews, cover art, tables of contents, excerpts, and any other kind of content that could help a person find out what a book could tell them. Plus it will have marvelous browsing categories, a recommendation service that makes personalized suggestions, and a search engine that corrects misspellings and suggests alternatives when a search comes up empty.

Patrons will visit this library any time of the day or night, from any place in the world with Web access. They can order any book they want from the catalog. Some of the books are immediately available from the community library; others take a little longer to arrive from other libraries. We never waste a patron's time. An e-mail notice tells them when we have their book. They can pick their books up at their local library, or have them delivered directly to home or office. Some fees involved.

Finally, this library has the potential to radically reduce the traditional costs of library operations, both by significantly reducing automation expenses, and by allowing libraries to restructure their physical collections to be more responsive to customer demand.

All sounds too good to be true, doesn't it?

But then a few short years ago who could have imagined that a bookstore could have 3 million volumes and no store to put them in?

Or that you could buy a book at 2 a.m. in the morning?

Or that a book catalog could learn to greet you by name?

Or that a small, upstart bookseller could grow to do more than a billion dollars of business in less than 3 years?

If they can do it, why can't we? And if we don't do it, who will?

Now if I could just figure out what to name this thing.

Mammoth — nah, reminds me of skiing. Humongous — nah, too cute. Gigantic — sounds like a grocery store. Everest — makes me think of bottled water. Amazon II — has a certain ring to it, doesn't it? I'll settle on a name tomorrow. Been too long on the road. Time to get some sleep. Tomorrow's going to be a very full day.

OCLC: "NO COMMENT"

EDITOR'S NOTE: Checking with one obvious potential player, we forwarded a copy of this article to executives at OCLC along with a list of interview questions. They found the piece very stimulating, but complex, and "decided to pass as to making any comment."

Hmm.

Author Steve Coffman pointed out that any number of alternative sources could start the process rolling. As Amazon.com started with Bowker's *Books in Print*, a forward-looking library

or library vendor could start with the Library of Congress' MARC listings or some major union catalogs. The trick, according to Coffman, is to start.

*Steve Coffman is Director of **FYI**, at the County of Los Angeles Public Library **(http://colapublib.org/fyi)**. His e-mail address is: **coffman@cerf.net***

• **Table of Contents**

• *Searcher* **Home Page**

Coffman, Steve (1999) Building Earth's largest library: driving into the future Searcher 7(3) 1999. At [24/03/01] Cohen, Suzanne et al. (2000) MyLibrary: personalized electronic services in the Cornell University Library 28. D-Lib Magazine 6(4). At [07/02/01] Corfield, Anthony, Dovey, Matthew and Tatham, Colin (2002) Z39.50 and XML: bridging the old and the new. At / [27/10/02] Cornford, James (2001) Integrating local resources Library Management 22 1/2 19-20 Cousins, Shirley (1999) Virtual OPACs versus union database: two models of union catalogue provision.