

Cites & Insights

Crawford at Large

Libraries • Policy • Technology • Media

Sponsored by YBP Library Services

Volume 5, Number 6: April 2005

ISSN 1534-0937

Walt Crawford

Bibs & Blather

Go Away—Not Now, But Soon

Have you planned a vacation this year? Great. If not, why not? It's been three years since I admonished readers to "get outta town!" (*Cites & Insights* 2:4). Then, as now, I know too many people treat vacations as disposable extras, niceties when nothing more important is happening. I don't believe that's true. Vacations are vital to healthy, balanced lives. Planning a vacation can be part of the fun, if you do it right.

Make It Real

Real vacations mean vacating—leaving home, leaving work behind, ideally leaving your technology behind as well. Taking a few days to get stuff done around the house (or lie around reading and taking walks) is great, but it's not what a vacation should be.

To me, a true vacation means:

- Being away for at least a week.
- Being somewhere and doing something that discourages thoughts of work.
- "Turning off": ignoring your blog and your aggregator, letting email stack up, setting aside IM. Ideally, you'll leave your notebooks, PDAs, and maybe (gasp) cell phones at home, although that may be too much to ask.

Follow Your Heart

Some people get the greatest pleasure from repetitive vacations—going the same place every year. I believe that's great as *part* of a vacation plan, but there's a lot of merit to travel and discovery. Maybe one week at your regular inn or ranch or amusement park or ski resort, and another week doing something new?

As I noted two years ago:

I don't believe there's a *Cites & Insights* reader who lives more than two hours from an area worth exploring, whether in the U.S. or elsewhere. Most of us fail to explore our extended back yards; maybe this is a year to be a traveler near home. Is there a "wine country" nearby? (You might be surprised!) State and national parks you never paid attention to? Historic towns—or, for that matter, the big city you've never approached as an outsider?

That's still true. If you live in any of the 50 United States, I *guarantee* there's a commercial winery somewhere in your state—even though some of them don't make wine from grapes, and others bring in grapes from other states (Alaska doesn't grow a whole lot of wine grapes, for example).

Inside This Issue

disContent Perspective: Print•a•bil•i•ty	2
The Library Stuff.....	6
Net Media Perspective: Google and Gorman.....	9
Trends & Quick Takes	16
The Good Stuff	20

I won't suggest what sort of vacation you should take. My wife and I have been exploring the world by cruise ship, as time and money permit, and we love it—even if we now occasionally revisit the same areas because they're so wonderful (for example French Polynesia, Alaska, and soon Costa Rica). But we've also enjoyed driving vacations and, at times, vacations connected to conferences. "Chicagoland" has many interesting areas in addition to the delights of Chicago itself, for example—and San Antonio in winter can be a great place to visit.

I'm delighted to correct one comment from 2002: "Sad to say, one of America's great neotraditional vacation possibilities is almost gone." That was the Delta Queen Steamboat Company and its three authentic steam-driven sternwheelers, cruising America's heart-

land rivers. The parent company was overextended and went into bankruptcy; as I wrote that essay, only the Delta Queen was still operating.

Fortunately, another company purchased the three Queens and the name itself, re-forming the Delta Queen Steamboat Company as an operating entity. All three boats are running again. We haven't been on them under the new management, but I can vouch for the charm and genuine hokey Americana of the Queens—and how interesting the heartland rivers can be. The one-week cruise from St. Louis to St. Paul (or vice-versa) includes great stops and a fascinating part of the Mississippi, including more than two dozen locks and dams. We loved it. You might even find the new “split week” American Queen vacations interesting: They combine a three or four night New Orleans roundtrip cruise on the American Queen, the grandest and newest of the Delta Queen boats, with three or four nights in New Orleans itself.

Plan a cruise. Plan a train trip (while you still can). Look into places of interest within a few hours of your home. You don't have to break the bank. You do have to break your daily habits and thought patterns. Enjoy the differences you'll find if you look for them (which does mean getting away from McDonald's and finding local color). You don't have to go to Nuku Hiva for a touch of the exotic (although we did love it). Paducah has its exotic side as well.

Get away. It will do you good.

Policy and Library Technology

That's the title of *Library Technology Reports* 41:2 (March/April 2005). I recommend it—but I would, since I wrote it. The chapters cover thinking in policy terms; the copyright spectrum; technology, privacy, confidentiality, and security; policy prerequisites and technology limitations; policy, technology, and the digital corpus; library policies and social policy issues; and sources and resources.

The issue draws heavily from the thinking I've done in the last four years of *Cites & Insights* (and before that), but it's not a rehash of these discussions. I've tried to put technology into a policy framework and consider the disparate ways that policy and technology interact. It's a coherent overall view that you won't find here and that I haven't seen elsewhere.

You can purchase the issue separately from ALA (visit www.techsource.ala.org), although it's admittedly not cheap at \$63 for the single issue. It's a peri-

odical issue, not a book, so I won't make more money if you all run out and buy it. I believe it's an important and worthwhile overview.

HTML Conversions as a Fundraiser?

Volumes 4 and 5 of *Cites & Insights* now have all the HTML stories they're likely to get. That leaves volumes 1 through 3, 2001 through 2003. Maybe there are stories in those first three volumes worth providing separately. Maybe not. They're a nuisance to process for HTML, a little more so than later editions—not a huge effort, but a nuisance. Maybe I'll get around to it. Maybe not.

Here's a challenge: If you would like to see stories from earlier issues made available in HTML form, *pay for them*. Not me, but some worthy cause.

Send a donation of at least \$100, preferably over and above what you'd normally donate, to one of the following:

- Freedom to Read Foundation
- Nature Conservancy
- American Civil Liberties Union
- Doctors without Borders
- World Wildlife Fund
- America's Second Harvest or one of the local Second Harvest agencies
- Habitat for Humanity

Send me email (waltcrawford@gmail.com) indicating that you've done so. I trust you (to some extent). You don't need to dedicate the donation in any way, and I don't require a receipt or proof that you've made the donation.

For each email I consider legitimate (mostly meaning it's from a real person, and only one per person), I'll do HTML stories for one issue of *Cites & Insights*, working backward chronologically from 3:14. I believe there are 41 eligible issues. Heck, for \$4,100 to a variety of causes most of which I directly support, I'll do a little work.

If 41 of you make reasonable-size donations, the whole run gets converted. If you don't, it might or might not. Hey, a little charity is good for you anyway.

disContent Perspective

Print·a·bil·i·ty

Here's a quote from an April 2, 2002 AP story as carried on Yahoo! News: “Ryan has not bee with wrongdoing.” Here's another, later in the same story:

“Falwell and others were behind the creation of false documents in the secretary of state’s offic justify pay raises...”

I’m not picking on the Associated Press. Read on screen, those two sentences make perfectly good sense—but in Yahoo!’s “printer-friendly” format, something’s missing. More than you see here, actually: on the printed page, half of the second “e” in “bee” and half of the “c” in “office” are missing.

By the time this column appears, Yahoo! may have discovered that its “printer-friendly” format is oxymoronic—but that will leave too many other Web sites that get in the way of the most natural thing a reader does with good Web content: print it.

Some DISCONTENT columns cover oddities, exaggerate to make a point, or deal with issues I find amusing. This is a case where I find the behavior of content providers self-defeating in incomprehensible ways. I’m bemused by the ways professionally-designed Web sites get in the way of printing content so that it can be read offline. Bemused isn’t exactly the right word. As a user, frustrated, annoyed, even “mad as heck” all come to mind.

Why Printing Matters

I can think of three reasons why anyone would want to print Web content:

- They want to read the content and it’s more than a few paragraphs long. The most optimistic claims I’ve seen are that people won’t read anything longer than 500 words online. As one content-oriented designer (at NUblog) puts it, “the only people who don’t print Web sites are those without printers.”
- What you say is worth repeating. People want to save it to cite elsewhere.
- What you say is valuable—interesting or lasting enough that people want to save it for future reference or rereading.

If your content is short and worthless, you can skip the rest of this column: you don’t need to worry about printing.

Worst-Case Scenarios

Yahoo! represents the worst case: offering a printer-friendly option that makes the text unreadable and ultimately ruins the content. I’ve run into a surprising number of other content-oriented sites that are nearly

as bad, obstructing effective printing in one or more of these ways:

- Running off the edge and having no printer-friendly option. That happens at Holt Uncensored, a text-oriented site about independent bookstores, publishing, and related topics. As with most such sites, it’s also hard to read on-screen unless you have a high-resolution display *and* turn off left-hand control panels. Just today, working on a copyright cluster for *Cites & Insights*, I wound up at digital media association, www.digimedia.org—and there goes the text, right off the edge of the page (far enough that I can’t make sense of the printed results).
- Dark backgrounds on printed versions, which not only waste toner or ink but also make reading difficult. More than one book-related site falls into this trap, even though the proprietors, of all people, should know better. The sites are hard to read on screen as well, so maybe these writers just don’t really want to be read.
- Light text and forced-small text that prints out that way. I know designers don’t like normal-sized text to mess up the site’s look, but does that justify expecting readers to read text that prints at less than nine points?

When I encounter these problems, I can be charitable and assume that nobody at the site has ever tried printing any of it out. Or I can be less charitable (and more like the average browser) and assume that the site proprietors just don’t care about readers.

Major Annoyances and Minor Peculiarities

A professional society puts the entire text of a book on the Web, one chapter per file, ready for printing—but there’s a wide black bar down the left side of every printed page and the pages are in sans, even though serif text is known to be much more readable in print form.

“Printer-friendly” versions show up with huge color ads inserted in the midst of the text, just in case we didn’t see the ad on the screen. There goes half a buck worth of ink if we’re using inkjet printers, which will certainly inspire me to purchase said product.

Some multipage articles offer printer formats, but only if you request print for each online page, one at a

time—even though the articles are ten pages or longer and require full reading to make sense.

Hello, Jakob Nielsen, usability guru: In your infinite wisdom, you must think we're all children. Your Alertbox forces us to read oversized sans type, and we get the same ugly, paper-wasting type on the printed page. (Just one more reason not to be too concerned about your "authoritative announcements.")

A beautifully-designed combination Webzine and blog prints out feature articles in clean, standard-size, *justified* serif type—but three-color (or gray and black on a laser printer) stripes across the top of each page obscure portions of the text.

Your site offers printer-friendly versions, and they work—but the resulting pages don't show where the copy came from. The URL is obscure (because it's a printable version) and you've left off clear source identification. That was a wonderful article I just reviewed a week later; too bad I'm not quite sure where it came from or when it originally appeared.

I could go on, and likely so could many readers. We all make little mistakes, but some of these also strike me as cases where the "content people" have never actually used the site or at least never actually printed from it.

Getting It Better; Getting It Right

At one point, the *Journal of Electronic Publishing* yielded printouts of its lengthy articles with light, hard-to-read text and some of the other problems noted above. There's still a narrow gray bar down the left margin, but these days the printed articles are otherwise clean, clear, and carefully identified. A fair number of other sites have also cleaned up their acts. Of the content-heavy sites I visit frequently, more do it well than do it badly—which makes the mistakes stand out all the more.

The NUblog article mentioned earlier offers a few hints for good printable pages; it's easy enough to find good advice elsewhere. I wouldn't trust advice from sites that don't yield good printable pages, but the fundamentals seem clear enough. Make the text (or printer-friendly version) monochrome (unless the color serves a specific purpose). Leave out the ads: we've seen them. Label pages clearly, with who you are, when the content appeared, and the original URL. Let the browser do text flow: turn off the special features that force long text lines.

For goodness sake, let body text be "normal" or "medium" size. And why not let the user's preferred typeface prevail for printed versions? If the user hasn't made a choice, the default's probably Times New Roman, which works very well on the printed page. And if the user *has* made a choice of a font he or she finds highly readable (and that is the point, isn't it?), he or she will appreciate having that choice honored.

Most of this boils down to "strip out the funny stuff." There are few things simpler in HTML than creating a printable page. Why make using the content everyone is so anxious to get up on to the Web and viewed by as many eyeballs as possible so difficult?

This DISCONTENT column appeared in *EContent* 25:7 (July 2002), pp. 40-41—exactly as it appears here, including significant editorial improvements by Michelle Manafy.

Blog Printability: Bringing the Story Forward

Readers with long memories may note that the most recent DISCONTENT reprint-with-postscript (in *Cites & Insights* 4:14) was the February 2002 column—and that I've normally included these columns in chronological order, skipping those that don't work well within *Cites & Insights*. I jumped to July 2002 because the printability problem hasn't gone away—and it may be getting worse in the blog world. Why does that matter? To quote and expand:

I can think of [several] reasons why anyone would want to print [Weblog] content:

- They want to read the content and it's more than a few paragraphs long [possibly including comments on your entry].
- What you say is worth repeating. People want to save it to cite elsewhere.
- What you say is valuable—interesting or lasting enough that people want to save it for future reference or rereading.
- They've been away from the blog for a while and would just as soon catch up in print form, reading a paper copy of recent entries.

Several of the weblogs I monitor *do* include essays more than a screen long. Some weblogs draw enough (and interesting enough) comments to make the entry plus comments worth printing, possibly running to several pages. And for secondary sources like me (as

opposed to other blogs), printing and saving is the only way blog material will be mentioned.

The Problems and the Triumphs

Problems with printing at text-oriented websites (that is, articles, papers, arguments, etc.) include the ones mentioned in the column. I haven't seen those problems that often in weblogs when prepared for printing or in text sites that use weblog-like tools for content management. I have seen a group of other problems, specifically the following:

- **Blogger's moving strip:** This phenomenon prints the body of a weblog in a narrow strip down the center of the page (sometimes with blog overhead such as archives and blogrolls on either side). Blogger's special trick: the strip starts moving to the right on each successive page, until part or all of the copy simply disappears off the right-hand edge of the paper. I've seen that happen as early as the third page, as late as the seventh page. It doesn't always happen at all.
- **One-page wonders:** Weblogs that stop after one page of the blog text. Period. Marking the text to "print selection" won't help. The *only* way I've found to print longer entries from these weblogs is to email the entries or to copy all the text into some program that knows how to print, such as Word. Most one-page wonders use Movable Type, but I think I've seen one Blogger one-page wonder as well, and one or two where I wasn't sure of the software.
- **The banner stands alone:** Weblogs that print the blog's banner or heading on a page that's otherwise blank—and then print one page of the blog's body and stop. In some cases, I've even seen a *blank* first page followed by a one-page wonder. This phenomenon seems to be a TypePad specialty, but I've seen examples from Movable Type and unknown software. Other weblogs using Movable Type and TypePad print the banner on an otherwise-blank page—but at least they let you print more than one page of postings.

At the other extreme, quite a few weblogs produce *cleaned* print versions: Printouts that omit weblog overhead and are clearly designed specifically for printing, using a separate stylesheet. These use paper

efficiently and are easy to read. Cleaned-for-printing weblogs are a WordPress specialty, although I've seen a few using other blogging software.

The Numbers

I didn't check eight million weblogs—I don't even claim that this is a representative sample. I checked all of the weblogs in my Bloglines list and most (but not all) of the other weblogs at LISFeeds.com. I also included seven "text-oriented" sites that I check separately, including two journalism magazines and a librarianship ejournal. In all, I looked at 177 websites.

Most websites clearly identify software, with four programs predominating:

- Blogger: 55 sites, 31% of the test.
- Movable Type: 39 sites, 23% of the test.
- WordPress: 24 sites, 14% of the test.
- TypePad: 12 sites, 7% of the test.

The other 47 sites (27%) either didn't identify the software or used programs such as slashcode, zope, LiveJournal, scoop, or IBlog.

Here's how I would judge the sites as letter grades—noting that this is *only* for printability, not for quality of content:

- A (ideal printability, typically "cleaned" for printing): 30 sites or 17%.
- B (good but not ideal—typically a strip wasting lots of paper): 87 sites or 49%.
- C or D (significantly flawed): 9 sites or 5%.
- F (impossible to print out the content in its entirety with readable results): 51 or 29%.

Nearly three out of ten sites tested were printer-hostile: That's an *awful* track record, particularly given that most of the sites here are related to librarianship or copyright. Don't people care about whether their words are read and retained?

Let's break that down by software:

- Blogger sites: one (2%) rated A, 23 (42%) rated F.
- Movable Type: two A (5%) and 11 F (28%).
- WordPress: 23 sites (96%) rated A; the other one was a solid B.
- Typepad: no A—and nine F (75%).
- Others: four A (9%) and eight F (17%).

Winners and Losers

I'm going to name the "A" and "F" sites—noting again that the grade *only* applies to printability! There are "F" sites that I like quite a bit; there may be "A" sites

that I wouldn't read on a bet. There is an order to each list, but it has nothing to do with excellence or awfulness. I'm not going to attempt to replicate the orthography and wordspacing of blog names.

Winners: All of these sites offered first-rate printability: Blog without a Library, Caveat Lector, Creative Librarian, Etc., Infomusings, Information Wants to be Free, Librarians Happen, Library Clips, Library Voice, Library Web Chic, Quædam Cuiusdam, Right-Wing Librarian, SciTech Library Question, Tangognat, Ten Thousand Year Blog, Furdlog, Blog of a Bookslut, Columbia Journalism Review, Web Pages that Suck, Bizgirl, Digitization Blog, Distance Education Library Services, Eng Lib, Hidden Peanuts, Kevin's Worklog, Library Lovers' LiveJournal, Library Monk, Library Grrrrrl, Linux Librarian, Netbib.

Losers: This list includes some of my favorite weblogs (and a lot of others)—but they sure do resist printing: A Wandering Eyre, Blog Driver's Waltz, C&I Updates, Canuck Librarian, Dave's Blog, Exploded Library, Icarus, Info Ediface, Infozo, ISBlogN, It's All Good, Librarian Avengers, Librarian in Black, Library Technology in Texas, LibraryLaw Blog, Library Technics, Rabid Librarian, Rick Librarian, Tame the Web, Tinfoil + Racoon, Twisted Librarian, Via Proni, What's New at OhioLink, Improbable Research, InfoThought, Kept-Up Academic Librarian, Mamamusings, Many-to-Many, Holt Uncensored, LIBRES articles, Online Journalism Review (blog and articles), Cog Sci Librarian, Collecting My Thoughts, Connie Crosby, Conversational Reading, Convivial Librarian, Distant Librarian, Drizzle, Feel-good Librarian, Lethal Librarian, Library Boy, Library Dust, Rambling Librarian, Schwagbag, Shush, Stephan Gallant Review, Teacher Librarian, Technogeekery for Librarians, Texadata, Unclassifiable Librarian.

I'd love to revisit the sites just mentioned in a few months and find the printability problems cleared up. It's clear that all blogging software *can* yield full printability; it just seems to be easier (or more likely the default behavior) with WordPress.

The Library Stuff

Block, Marylaine, "Libraries: The original 'long tail,'" *Ex Libris* 239 (February 11, 2005). marylaine.com/exlibris/xlib239.html

I've mentioned the "long tail" before—and the extent to which I believe it's a fairly typical *Wired Magazine* situation: An editor grabs a long-standing cultural phenomenon, gives it a cute name, generalizes, and claims it's something New and Special. The concept that most people appreciate and buy (or consume) media and other options far beyond the best-seller list should be familiar to libraries. It's certainly familiar to good bookstores, magazine publishers, book publishers, record companies, and Netflix. Calling it "the long tail" gives Chris Anderson a wonderful new discovery and most likely a book that will be one of those irrelevant best-sellers. Oh that's right: Anderson says the *Internet* makes the long tail feasible—which is largely nonsense but gives the concept that digital aura of greatness and newness.

Here's Block's comment after quoting a typically breathless Anderson paragraph:

I submit that the only thing new about the long tail is that because of the internet, the commercial world is just now discovering it. Libraries have been in the long tail business for centuries.

I disagree in part. Magazine publishers, very much part of the "commercial world," *rely* on the "long tail": 99% of the quarter-million magazines published in the U.S. reach tiny minorities of the reading public (less than 1%, mostly *much* less than 1%).

Block goes on to quote from an email conversation between her and Anderson (who apparently knows almost nothing about libraries, another consistent *Wired* trait).

That said, **read this column**. Anderson may be as tired as the rest of *Wired*, but Block has good things to say, particularly about the importance of libraries maintaining a commitment to deep collections (call it the "long tail" if you must) along with improving marketing savvy.

Hodgson, Cynthia, "RFID in libraries: Are we ready?" *Information Standards Quarterly* 16:4 (October 2004): 1-5.

I'm delighted that *ISQ* is still around, going strong after 16 years. I was the first editor of this NISO publication, which replaced *Voice of Z39* in 1989. I edited and desktop-published 12 issues (three volumes) before turning the editorship over to Pat Ensor.

The lead essay in this issue is **well worth reading** as a good discussion of RFID's strengths and weaknesses. Hodgson lists nine technological strengths (from ease of reading through reliability and life ex-

pectancy) and seven weaknesses (including use of proprietary technology, cost, and possibility of signal interception). After discussing library applications, Hodgson goes on to consider privacy concerns in one of the most levelheaded discussions I've seen. She notes standards development in the RFID area and concludes that some libraries may find RFID implementation timely while others may choose to wait for lower costs or to resolve privacy and other issues.

Smith, Kathlin, "The value of library as place," *CLIR issues* 43 (January/February 2005). www.clir.org/pubs/issues/issues43.html

This four-page article is an introduction to a forthcoming CLIR publication. The article begins with a provocative paragraph and two seminal questions:

Google's recent announcement that it will collaborate with several major research libraries to digitize portions of their collections has brought the promise of desktop access to large research collections closer to reality. At the same time, it rekindled discussion of a question that emerged as soon as the potential of the Web became apparent: Do people still need the physical library? Most people agree that we will continue to require physical repositories of books and scholarly materials. Yet the wealth of high-quality information that can be accessed now, and the promise of more to come, challenges the library's traditional reason for being: to serve as a repository of information and to make that information available to users.

What is the role of a library when users can obtain information from any location? And what does this role change mean for the creation and design of library space?

"Repository of information" is a sad version of a library's role as collector, organizer, and repository of the cultural record. There's a lot more to any good library than "information." Indeed, the equation of library collections with "information" is one of the oversimplifications that encourages futurists to suggest that the physical collection can disappear when (almost) everyone has (some) access to (loads of stuff that may include) information on the web. Very little of what I borrow from public libraries can be classified as "information," and most of my information doesn't come from library collections—and I regard physical public and academic libraries as fundamental to a civilized society.

Architect Geoffrey Freeman discusses the reintegration of teaching spaces into academic libraries—a reintegration that can be positive, as long as it's not

essentially impoverishing a library by snatching its space. Scott Bennett, Yale librarian emeritus, argues that libraries need to understand more about how students learn in order to design space that supports those needs. One essay provides examples of how institutions are exploring a library role as "laboratory for the humanist and social scientist," while another considers libraries that serve both researchers and the public. This should be a worthwhile volume.

Stanescu, Andreas, "Assessing the durability of formats in a digital preservation environment," *D-Lib Magazine* 10:11 (November 2004). www.dlib.org/november04/stanescu/11stanescu.html

"There are two necessary components in any measuring system—the units of measurement themselves and the process of applying those units of measurement." This article offers a methodology for measuring the preservation durability of digital formats—something that Stanescu concedes isn't as clearly definable as a meter. He hopes that INFORM, the methodology described in the paper, "will be the first step towards creating a useful definition for preservation durability."

If you've been paying attention, you know that long-term digital preservation involves not only the longevity of digital media themselves but also the *understandability* of what's stored on the digital media. "Given the speed at which formats come and go, how can modern librarians and archivists identify those formats most apt to survive the passage of time?"

INFORM, Investigation of Formats based on Risk Management, "attempts to discover specific threats to preservation and measure their possible impact on preservation decisions." Stanescu mentions six classes of risk: risks arising from the digital object format, software, hardware, associated organizations, the digital archive itself, and migration plans for preservation. The methodology yields a "risk exposure" for the format that may change over time as risk is reassessed.

This paper is a fairly brief introduction to a complex topic. As such, it's certainly **worth reading** if you expect to be involved in preservation decisions.

Stephens, Michael, "The *balanced* librarian," *Tame the web*, January 28, 2005. www.tamethe-web.com

Michael Stephens calls it "unplugging." I didn't give it a specific name—other than contemplation—in the March 2003 "The Crawford Files," "The century's

most vital technological device” (*American Libraries* 34:3, p. 84). The idea’s the same: Unplugging. Taking breaks. Stephens works out 4 days a week “plugging in only to my iPod.” I’d go further: You need time when you’re not plugged in to *anything*.

Stephens also discusses the need to pick and choose, to select a few interests to follow to avoid drowning in too much everything. He’s right—and it’s good to hear this from someone as committed to All Things Technological as Stephens. If you don’t normally follow *Tame the web*, go track down the January archives and read this post. For that matter, go back and read my column. (And see the opening of BIBS & BLATHER, which was *supposed* to be a PERSPECTIVE in the Spring issue before...well, before things got too crowded and busy.)

Tennant, Roy, “A bibliographic metadata infrastructure for the twenty-first century,” *Library Hi Tech* 22:2 (2004): 175-81.

When Roy Tennant wrote “MARC must die” in *Library Journal* (October 15, 2002), I gave him grief for it (at least privately). Not because my first book was about MARC and because I’ve been using the format for mumbletysome years (at *least* 30). Not because I think MARC is perfect. My problem with Roy’s columns was that they struck me as oversimplifying the situation and underestimating the sheer difficulty of replacing MARC with anything else, at least without losing much of MARC’s specificity and granularity. But he was trying to bundle a bunch of ideas into two very short columns—and he was trying to get people interested in a relatively arcane topic.

This article suffers from neither problem. I **recomm**end reading it if you have access to *Library Hi Tech* and if you’re interested in a workable future for bibliographic retrieval that goes beyond full-text keyword searching. Tennant here does *not* say that MARC must die. What he says is that we need (and are building) an infrastructure that can accommodate a variety of metadata representations, some much simpler than MARC, some adding even more granularity to MARC (it’s needed in several areas such as personal names as well as 773\$g)—and, for an enormous quantity of valuable bibliographic information from the past and going into the future, MARC itself:

What must die [are] not MARC and AACR2 specifically, despite their clear problems, but our exclusive reliance upon those components as the only requirements for library metadata.... We must...assimilate

MARC into a broader, richer, more diverse set of tools, standards, and protocols.

It’s hard to disagree with that assessment. MARC as a set of metadata elements isn’t going away now or anytime soon. Z39.2, the formal record structure used to transmit MARC records between systems, probably isn’t going away either: It’s far more compact and self-contained than most metadata structures, and I know from experience that you only need about a dozen lines of code in a high-level programming language to be able to decode MARC at will. But MARC isn’t enough, and hasn’t been for a long time now. Yes, you can shoehorn lots of other things into Z39.2 and MARC records—but that’s force-fitting, and it shouldn’t be necessary.

This article discusses reasons the library community needs more than MARC itself—among them granularity, hierarchy, and the need to accommodate both very simple and very complex metadata—and some ways to deal with those needs. It’s happening, as Roy points out, and it’s being done in the same cooperative manner as the MARC formats: a collaborative effort involving RLG, OCLC and other institutions, coordinated by the Library of Congress.

OCLC’s new WorldCat will store records in an XML format that can incorporate MARC21 as well as other metadata systems, at least over time. RLG’s new database environment stores records in an XML format that can incorporate MARC21 as well as other metadata systems, at least over time. We’re all becoming familiar with crosswalks; we’re all struggling with issues of merging records and determining how the FRBR concept will play out in practice.

MARC isn’t dying. It’s becoming part of a richer community of bibliographic metadata standards. This is a good thing—an essential thing for anyone who believes there’s more to searching than tossing words against a sea of full text, for anyone who believes cataloging is meaningful, and for anyone who understands that the world has moved beyond a single universal format.

Wilson, Thomas C., “The origins of *TER*: Ten years after,” *Technology Electronic Reviews* 12:1 (February 2005). www.lita.org/ala/lita/litapublications/ter/volume12no1.htm

I rarely mention *TER*, not because it isn’t worthwhile but because I rarely mention book reviews—and that’s what *TER* is all about. It’s probably LITA’s first electronic publication and certainly the most suc-

successful one to date: Founded June 2, 1994, it's done quite a bit since then, as Tom's editorial shows. **Worth reading** if you care about the history of ejournals in librarianship.

I'm amused that Tom gives a formal name to my informal measure of longevity for ejournals: that an ejournal that's still publishing and has a minimum of six years' publishing record can be considered a lasting title. He calls it the Crawford Test. I'll stand by six years as a reasonable measure for significance: Even if it later fails, any journal, e- or print, that publishes regularly for six years has done something right.

Net Media Perspective

Google and Gorman

Whether you call it Google Library or just part of Google Print, Google's massive scanning-and-OCR project has just begun. My "non-comment" in January was all I intended to say—but it turns out I had said a little more in an informal conversation with an *American Libraries* editor. This story probably won't go away for years and deserves more attention.

Google Print: Prototypical Reactions

In some ways, reactions to Google's announcement of its deal with five major libraries were almost prototypical: "Here's how techies, librarians, and others respond to some big new project that could impinge on their lives." Consider some stages—summarized without naming specific names, and if you think that means these are straw men, you've been hiding in a cave somewhere:

- Google makes an announcement that *could* mean that *eventually* the text of more than ten million books *may* be searchable on Google, either enmeshed with billions of other documents or separable as a book-oriented search service. Google does not promise to make these books available in "ebook" form and doesn't really provide many details.
- Some doom-crying librarians say it's the beginning of the end for libraries: If all the books are on Google, why would anyone need a library?
- Some exultant techies say it's the beginning of the end for print books and traditional catalogs: After all, what could be better than full-

text searching and a universal collection of ebooks? And it's all free!

- Some librarians (and others) denounce Google, the project, or both, for a range of reasons—some sensible, some exaggerated, some over the top. We hear accusations that the Google project means libraries are being commercialized and users will lose privacy. We read that this plan furthers the Anglo-American hegemony of the web and is an insult to other cultures, specifically France. We're told that Google is an incompetent search engine—as opposed to the *true* statement that fielded searching works better for many forms of scholarship than Google's full-text searching. Some critics assume the books will be "disbound" (that is, their spines will be cut off) and discarded after scanning, because that's the way a lot of fast, cheap scanning has been done in the past. Others assert the scans must be extremely low resolution. (These statements appear *after* statements to the contrary from program participants have appeared.) We hear charges that Google will be *the* gatekeeper for all knowledge, perhaps aided by Google's ambitious plans.
- A few "librarians" publish columns that seem to imply that this project will make physical libraries, or at least their book collections, redundant—and *celebrate* that "fact" because it will free librarians to become something better. Such as searchers, I guess.
- Some publishers or publisher groups complain that scanning books and making them searchable online violates publisher copyright—even though Google has said that nothing more than a tiny excerpt of copyright material would be available and that there will be links to ways to locate and buy books, and even though *sensible* publishers know that online access has (so far) consistently increased sales of print books.
- Some librarians and others offer thoughtful comments falling somewhere between "the sky is falling" and "the skies are opening," recognizing both the potential and the limitations of the project—and recognizing that it's barely begun.

- Meanwhile, the project proceeds along a course that seems to offer no likelihood of undermining libraries or publishers and every likelihood of increasing library use and book sales. Libraries get their books back presumably unharmed (I'm certain these libraries are monitoring that issue) with the promise of digital copies they can use for other purposes.

Google Print *should* be a win-win-win-win situation for Google, libraries, readers, and publishers. I still doubt it will be as successful as some hope, and I am among those who believe that Google's results aren't as focused now as they were when the index had a mere billion or two documents. I don't believe full-text keyword searching with "relevance" ranking can or should fully replace fielded searching of bibliographic records with authority control—but I do believe that full-text keyword searching is another powerful tool for formal and informal scholarship and learning, and that Google does it at least as well as any search engine.

The "page 140" scenario

Michael Gorman is among those who I believe has made exaggerated negative statements about Google and Google Print. One of those statements keys off a paragraph in *Future Libraries: Dreams, Madness & Reality*:

One writer asserted that a good way to choose a book in order to learn a new subject is to open it to page 14, page 54, and page 140. If the reader cannot make sense of each of these random pages, the book is too complex. The proponent of this theory has spectacularly missed the function of a good linear text[s] intended to impart knowledge. By the time a reader has read pages 1 to 140, he or she should understand the subject much better than when on page 14. Indeed, if a book is designed to inform on a subject that is new to the reader, that reader has every right to be suspicious if all of, say, page 154 is instantly accessible. What is the point of reading the text in that case?

I'll stand by that paragraph. Where things go awry is when that logic is used to suggest that locating words of interest on page 154 is of no use. Quite the contrary. Let's say I know nothing about the connection between quantum theory and superstrings—heck, I don't know what the words mean. Let's say that, in 2014, I go to Google Print and search "quantum theory" "superstring"—and get back a paragraph on page 154 of a book on physics.

Chances are good that the page won't make a lot of sense to me on its own. But I will get two things:

- A citation for one book that discusses these exotic terms. With that citation, I may be able to click through to a service that will show me a call number, which may lead me to related books, one of which might be at a level I can understand. (I may get a direct link to a source for the book itself, for that matter.)
- Maybe enough text to provide some indication of whether the *level* of the book will suit me. Maybe not—but there's at least a chance.

Sometimes, of course, a paragraph may give me the hints I need to proceed toward more effective research. Sometimes, the paragraph may strike me as complete gibberish. Chances are, if the project succeeds, I'll get more than one paragraph from more than one book, including not only books digitized from libraries but also current books being provided by publishers. Somewhere within those results, there may be help to move me on my way: Not necessarily the answer, but a starting point.

Sample reactions and comments

As related in a February 15, 2005 post on *LibraryLaw blog*, Ernest Miller had an interesting response to publishers' assertions that the Google project infringed on copyright: The results of the project "could be seen as a really efficient index"—and in a similar case that was found to be fair use. "Looking at GoogleLibrary as an index, displaying only snippets, I think a strong argument can be made for fair use..."

A few days later, Mary Minow at *LibraryLaw blog* wondered whether libraries were considering the commercialization aspect of "turning over massive collections for digitization by a commercial player." She's right—"we must ask the right questions in the early stages."

Steve Johnson offered a brief overview in "How Google will scan the world, 1 book at a time," February 25, 2005 in the *Chicago Tribune*. He suggests Google has ubiquity as a goal—"and as Bill Gates has learned, when you become unavoidable, you also become resented." Johnson quotes Sidney Verba (Harvard) on Google's plans and the fact that "the libraries will also get their own copies of their texts turned binary." Johnson also quotes Michigan's comment that its own digitization project—one of the best in the country—would take "more than a thousand years" to digitize the 7-million-volume collection at the current

rate; Google expects to do it in “a matter of years.” Finally, there’s Verba on the “threat” that Google Print poses to libraries: “The nice thing about this project is that it’s a kind of, ‘If you can’t beat ‘em, join ‘em.’ People will go to Google, and they will find books, and they will then go to the library and get the books.”

Barbara Fister’s March 2005 piece in *Library Issues* notes some of the controversies surrounding the project. She says Google’s scanning of post-1922 books is “almost certainly a violation of the law” and offers this trashing of the entire public domain:

Three of the five libraries participating are only digitizing books clearly in the public domain. That means if students do encounter books through a Google search, chances are they will be so out of date as to be useless. Even classics will only appear in inferior editions, without the benefit of contemporary introductions and annotations, careful editing or the inclusion of recently-discovered material.

Fister goes on to offer thoughtful consideration of some of the real issues, noting and discarding the extreme views (the wonder of universal free access to all books, and the horror of drowning in a sea of undifferentiated information). I’m not sure I’d agree that full-text searching is “nearly useless for the novice researcher,” but she’s right that subject cataloging, classification, and citation networks offer surer and generally better ways to find solid research—at least when they work. She also dismisses the idea that cyberspace would substitute for physical libraries. “Google may have market dominance among Internet search engines, but it isn’t out to replace libraries.” True—and Google’s clear about that.

Dorothea Salo offered a valid warning about some Google Print hype in a March 11, 2005 *Caveat lector* posting. She notes that scanning and OCR don’t turn a print book into a usable digital object, at least not if the book has interesting characteristics. Salo goes so far as to say, “Scanning/OCR is the easy part... It’s everything else that’s hard.” She cites some obstacles to creating fully usable “book replacements” without lots of hard work. I believe her posting (and some other non-hysterical but critical articles and posts) are important cautions for those who believe Google’s going to make 14 million ebooks available—but I’ve heard nothing to suggest that Google has such an intention. If the intention is to provide medium-quality searchable text, screen-quality (but non-printable) pages for old books and snippets for newer books, and links to ways to find the actual books (along with, presuma-

bly, contextual ads along the side), scanning and decent-quality OCR may be just what Google needs.

I do know that Google is neither a replacement for catalogs and professional indexing, nor the only good web index, nor yet again a worthless pile of junk. I do know that Google is a stock corporation, which will tend to interfere with its laudable long-range planning—but it may be sufficiently closely held that this doesn’t pose a problem (and it’s profitable enough that Google Print may constitute a worthwhile long-term experiment). I do know that many librarians and technogeeks will happily overreact to any new announcement in a variety of ways, sometimes too cynically, sometimes too optimistically—sometimes both at once.

Michael Gorman called this and similar projects “expensive exercises in futility” and “a solution in search of a problem” in a Los Angeles *Times* op-ed—and later, in the March 2005 *American Libraries*, said, “Any user of Google knows that it is pathetic as an information-retrieval system—utterly lacking both recall and precision, the essential criteria for efficiency in such systems.”

I’m not saying Gorman got it all wrong. He didn’t. I agree that calling Google “the gatekeeper to the world’s knowledge” is “a combination of hype and hubris.” I agree that, *for many purposes*, free-text searching is “inherently inferior to controlled-vocabulary systems.” But Gorman goes overboard, as already noted—and, I think, goes too far in saying that the project will do nothing to assist digital preservation. (Since copies of the scans are being given to the libraries, and if those copies are open to suitable uses, that’s a step in the right direction—even if Google Print itself does little to preserve the copies.)

Quite a few bloggers and others on the internet commented on Gorman’s *LA Times* piece. Some of them did so intemperately, favoring passion over literacy and thought. Some of them blasted Gorman, sometimes far beyond what his piece deserved.

Gorman and the Blog People

Using my standard practice, I should be calling Michael Gorman “Michael.” I’ve met him on several occasions (after all, we did coauthor a book). But Michael Gorman is incoming president of the American Library Association, provided profound service to the library profession as editor of the *Anglo-American Cataloging Rules, Second Edition* and succeeding works,

and was a leader in introducing technology within academic libraries for many years. So Gorman it is.

Gorman is at a level where he can get an op-ed published fairly easily, and *Library Journal* just loves controversy. The February 15, 2005 *Library Journal* featured a “BackTalk” entitled “Revenge of the blog people!”—and that’s when things started getting crazy.

I don’t see how you can be a public university library director without growing a thick skin. I *certainly* don’t see how you can participate in ALA-level politics, particularly running for president (twice) without learning to ignore most insults. Part of developing a thick skin—the part I’ve never been good at—is remembering that “ignore” means “don’t respond.”

Gorman responded. Boy, did he respond. Here’s the first paragraph of his response:

A blog is a species of interactive electronic diary by means of which the unpublishable, untrammelled by editors or the rules of grammar, can communicate their thoughts via the web. (Though it sounds like something you would find stuck in a drain, the ugly neologism *blog* is a contraction of “web log.”) Until recently, I had not spent much time thinking about blogs or Blog People.

I’d guess some flames were ignited and posted before people even *read* the second paragraph, given the intemperance and overgeneralization of that lead sentence. But wait! There’s more! He summarized his op-ed in no uncertain terms, saying of Google that “it gives the searchers its heaps of irrelevance in nano-seconds” and going on, “rubbish is rubbish, no matter how speedily it is delivered.” Later: “It turns out that the Blog People (or their subclass who are interested in computers and the glorification of information) have a fanatical belief in the transforming power of digitization and a consequent horror of, and contempt for, heretics who do not share that belief.”

It is obvious that the Blog People read what they want to read rather than what is in front of them and judge me to be wrong on the basis of what they think rather than what I actually wrote. Given the quality of the writing in the blogs I have seen, I doubt that many of the Blog People are in the habit of sustained reading of complex texts. It is entirely possible that their intellectual needs are met by an accumulation of random facts and paragraphs. In that case, their rejection of my view is quite understandable.

If Michael Gorman had interspersed a few “many” and “some” and “the more intemperate” qualifiers in the op-ed, I would only say that his comments about

Google are somewhat simplistic. I agree with Gorman that such comments as “Michael Gorman is an idiot” and claims that he’s “antidigital” or a “Luddite” are themselves overstated—and typical of the reasons I stay away from /. and, usually, Kuro5hin. I can empathize with the claim that some bloggers “read what they want to read rather than what is in front of them,” since I still feel I’ve been a victim of that tendency. But, but...there’s way too much generalization here, and that first paragraph is such a torch that one can hardly be surprised when flames resulted.

Not all flames

I read many intemperate responses to Gorman’s intemperate op-ed, even though I carefully avoided particularly flame-prone parts of the web. I was astonished by the number of supposed librarians who dismissed Gorman’s achievements and career on the basis of this one piece of writing, labeling him with the terms above and worse. I was saddened by the number of badly written, worse thought out comments that seemed to justify the worst Gorman could say about blogs and bloggers. A few days into the s**tstorm, Gorman characterized the op-ed as “satire”—but most people didn’t read it that way, and neither do I. (At best, it’s unusually hamhanded satire for an essayist of Gorman’s skills.)

There were also quite a few thoughtful responses, the kind I’d expect from many library bloggers and list participants. Brief notes on a few of those, taken chronologically as usual:

- Rochelle at *Tinfoil + raccoon* (February 24) agreed, “Plenty of blogs are pure crap, as are plenty of print publications... It shouldn’t be a case of bloggers v. traditionally published writers. Or Google v. libraries. Or book v. ebook. Or *World Book* v. Wikipedia. It should be a case of quality and authority v. crap. I’d rather teach those critical skills to people so that they can figure it out for themselves, rather than limit sources of information.”
- George Needham, part of the *It’s all good* team (three OCLC bloggers), wrote a wonderful brief satire on Gorman’s attitude toward Google Print in his February 25 post, “Revenge of the codex people.” I can’t possibly summarize this piece of carefully planned (and, I’d guess, carefully edited) work; go to scanblog.blogspot.com and find the February 25, 2005 archives. If you haven’t already read

it, **do**: it is indeed “all good.” A March 3 followup notes “what a blast I’ve had” reading comments and email—and noting *why* blogs are “not only fun, they’re important. They allow us to have a dialogue with people with whom we might never have a chance to interact in person. We actually do learn things, skeptics notwithstanding, and learn more about our colleagues.... My main concern is that I see the world being polarized in yet another way, among those who think electronic resources are a panacea and those who think they are a plague. We need to create and inhabit an ‘and’ world, not an ‘or’ world. It doesn’t have to be print *or* digital, it can be print *and* digital.” Ah, George, it does my heart good to see you using my long-standing motto, even if indirectly.

- Michael Stephens of *Tame the web* offered “An open (yet personal) letter to Michael Gorman” on February 25. He responds to one of the BackTalk subheads, “Who are the Blog People,” by describing himself as a librarian and blogger: “I love libraries—especially the public library. I’ve been with the St. Joseph County Public Library in South Bend, IN for almost 14 years... I have worked hard to improve services to our users via my particular passions: staff and public technology training, using new technologies to meet user needs and the juncture of tech, people and libraries.” As part of a generally eloquent commentary, Stephens responds to Gorman’s general attack on the quality of blog writing and blogger reading:

I am a doctoral student at the University of North Texas in the inter-disciplinary Information Science program and I have spent quite a bit of time reading scholarly works and professional articles. Since June I have been gathering information on online communities and the interactions of people within them. I have written literature reviews, proposals for research, critiques of articles and multiple blog posts on my experiences and thoughts about libraries and librarians. The community of practice I have seen spring up within the LIS Blogosphere has inspired me to participate, write better and seek opportunities to show librarians how such a simple thing as Weblog software can ease the dissemination of information and generate knowledge.

- Blake Carver only has a “blog lite” (his *LISNews* journal), unless you want to consider *LISNews* itself to be his überblog, but he contributed an eloquent post—partly in Gorman’s defense—at Web4Lib on February 25, as part of a heated thread on that list. “Blogs are indeed often unpalatable, and always untrammelled by editors. While he meant this to be insulting (or was it satirical?), us bloggers should be able to admit these are legitimate complaints. Though at the same time these are some of our greatest strengths... Most blogs are not great works of literary history, most bloggers are not great writers, but that’s not the point... The funniest part of the entire article was just how much this article was [like] a post I’d read on any day [at] *LISNews* or any other blog. A post that would probably get moderated as flamebait...”
- LITA-L had a similarly heated thread, including some comments from people as dismissive of bloggers as Gorman himself. Fred Stoss agreed that, paraphrasing slightly, those who can, publish; those who can’t blog. Leo Robert Klein responded, in part, “I think this doesn’t take into account a great number of Blogs where the writers *do* publish—and publish on a regular basis.” Brenda Battleson—a colleague of Fred Stoss—noted blogs as “a means of getting ideas ‘out there’” and offered references to correct those who think blogs are for those who can’t get published. She noted uses of blogs at her university and cited some particularly worthwhile blogs.
- I noted that most of what I make available for public consumption (specifically, this here semi-literate journal) bypasses editorial control and traditional publishing, “putting me pretty squarely in that ignorant semi-literate group of folks with nothing worthwhile to say.” Steven Cohen cited that and my serious consideration of starting up a weblog as “one of the neatest reactions to the Gorman attack on bloggers.” I wouldn’t go that far, but there’s little doubt that Michael Gorman’s attack served as a tipping point in the creation of *Walt at random* (<http://walt.lichost.org>), which is a weblog. I’m not a “blog person.” That’s a separate issue. I am, as of April 1, a widely

published writer who also produces a weblog. Make of that what you will.

- Andrea at *LibraryTechtonics* posted a fairly long and clearly well thought out commentary on February 27, 2005. She explicitly does *not* dispute what Gorman said about Google in the original *LA Times* piece: “honestly, as a technologist *and* a librarian, I agree with just about everything said about Google not being the über-answer.” She regards the *LJ* piece, however, as “yet another misunderstanding about blogs, blogging, and bloggers, but significantly more devastating coming from the publicly perceived leader of the library community.” She goes on to define blogging clearly—“A blog is a technology that allows people to publish to the web quickly and easily. Blogging is the act of publishing to a blog”—and to note that those who hold blogging to be a “philosophy, a way of life, a moral imperative” are a subculture, [probably] not the majority of bloggers. Andrea (yes, we’ve met—a considerable pleasure!) calls herself “a librarian with a blog,” not a journalist (on the blog). “I was an English major, I write well, I mind my grammar, I check my facts, I ask before I quote or paraphrase.” That’s her practice—more careful than mine and those of some bloggers, but also more careful than those of some journalists (which she explicitly says she is not). She goes on to offer several thoughtfully stated reasons she’s unhappy with the *LJ* piece. I suggest you **read it yourself** at www.librarytechtionics.info.
- Marylaine Block devoted *Ex Libris* 242 (March 4, 2005) to the controversy, titling the piece “Family feud.” Block likes Gorman and voted for him for president. “That doesn’t mean I will support him when he makes sweeping, defamatory comments on an entire class of people he doesn’t appear to know much about.” She notes that some of librarianship’s most respected writers are producing weblogs “and library publishers are finding some of their hottest new writing talent by reading library blogs.” She lists a few of the leading library weblogs and quotes her husband, “one of the few people in the world who could equally love an exquisite rare beef roast and

the cafeteria version of it. The trick, he said, was to regard them as two entirely separate dishes.” I might be like her husband in some ways—I respect a really good cheeseburger much as I respect a really good steak *au poivre*—and Block goes on to explain why the “two entirely separate dishes” of traditional published literature and weblogs have separate strengths.

- I was charmed by *Valisblog*’s March 8 “Very belatedly: GormanGate.” Simon (who I *haven’t* met, as far as I know, but he signs his posts with one name) is another blogger who, like Andrea of *LibraryTechtonics*, agrees with much of what Gorman originally said about Google. “Google is a great tool for some purposes. It is perhaps not such a great tool if you are using it to locate book-length scholarly materials... Because Google searches on the full text of documents, you run a fairly high risk of locating irrelevant information that, by chance, contains your chosen search terms.” Simon goes on to call the *LJ* piece “a fairly stupid, poorly written response, that implied that a whole class of people were ignorant and incapable of reading whole texts.” Still, Simon was surprised at the sheer vehemence of much of the reaction. “A lot of people seem to be getting awfully defensive...” (I just realized Simon indirectly complimented me, if only as a co-author. Thanks.)

So what does it boil down to? Gorman attacks Google (in certain contexts). Bloggers attack Gorman. Gorman attacks bloggers (in general). Other bloggers get even more upset. Seems to me like there’s been a fair bit of over-reaction, all the way through this episode. Deep breaths, people. Deep breaths. Even *if* Michael Gorman thinks you’re an idiot with no attention span, does it really matter?

- Still later (March 16), Jane of *A wandering eyre* used the whole fiasco as the jumping-off point for an essay on the nature of leadership. It’s a solid, thoughtful essay, one that (again) deserves **reading on its own**: wanderingeyre.blogspot.com.
- Rory Litwin also felt compelled to consider how *Library Juice* compares to a blog (in *Library Juice* 8:6 at www.libr.org/Juice/issues/vol8/LJ_8.6.html), after noting that he’s also been critical of the “blogging craze” but that

he regards blogging as “more akin to casual conversation taken into the realm of the printed word” and doesn’t consider the “blog people” to be a cultural threat:

Potentially more of a threat to serious publishing (and I’m being half-serious here) are publications like *Library Juice*, which, though not blogs, borrow some of the attributes of blogs while still making a claim to be a part of the world of publishing.

He cites reasons *Library Juice* is a traditional serial: fairly regular periodicity, ISSN, an editor “who vets contributed articles for quality and fittingness,” mostly *not* written by the editor, an editorial perspective and some consistency, traditionally citable, and with longer and more “traditional” articles than in a typical weblog.

He also notes the reasons that *Library Juice* (and *Cites & Insights*) are sometimes called blogs: Freely accessible via the web; an RSS feed (or at least part of one); “very much a DIY project”; a more casual editorial process than most established print editors; and functionally “part of the blogosphere.” Litwin then admits to being “one of the blog people.” He understands the negative implications but says he’s “going to keep going as I’ve been going.” And he clarifies his “craze” problem: his sense that blogs “have become the default format for any new website, regardless of the appropriateness of a centrally chronological organizing principle.” There’s more, **worth reading**.

My Own Take

Gorman admits that he “had not spent much time thinking about blogs.” There’s nothing wrong with that. For all the claims of certain hotshots, blogs have *not* replaced journalism, scholarship, or traditional media. Blogs add more than they replace, and what they add doesn’t work for everyone.

Consider Sturgeon’s Law: 90% of everything is crap. Theodore Sturgeon was responding to mainstream critics who had dismissed science fiction with the comment that it’s mostly crap. I’ve always used Vivaldi as an example to suggest that 90% of Baroque compositions were probably crap as well (even though most of that 90% has presumably long since vanished). Sturgeon, however, was commenting on *published* material, and probably half of the material submitted to traditional publishing is rejected. Add that in, and you could suggest that 95% of all writing, or music, or art, or whatever is likely to be crap—at

least if you define “crap” as being of no lasting importance or interest outside of a small circle.

There are eight million weblogs as of one count. If there are 400,000 weblogs of lasting importance or interest outside of the blogger’s small circle, that’s remarkable. I doubt the number is anywhere near that high. I would suggest the number is considerably north of zero, probably in the tens of thousands.

Many bloggers write badly. The medium lends itself to “zeroth-draft” writing, where you type an instant comment without thinking it through carefully. Many bloggers think badly, oversimplify, exaggerate, and have incredibly thin skins. Many bloggers equate skepticism with bias and doubting with opposition. Many bloggers think digital is automatically better and have no interest in reading anything longer than a screen of text.

I say “many” with some assurance because “many” can mean anything from a hundred to a million. I’d guess each of the nasty comments in the preceding paragraph applies to at least a thousand bloggers. Taken as a whole, I’d guess they apply to a minority of bloggers, possibly a small minority—but that minority is still a whole bunch of people.

Many bloggers write well, some superbly. Many bloggers edit what they write before they post. I suspect more than a few show their essays to others for editorial review before they publish them. Many bloggers think deep thoughts, clarify, explicate, and stand up for what they believe while accepting criticism and understanding that other people may think differently. Many bloggers understand the difference between skepticism and bias or opposition, and understand that you can regard something as worthwhile while criticizing its flaws. Many bloggers believe in the virtues of multiplicity and see digital tools as opening new avenues without closing the existing possibilities.

Within the library community, I believe the percentage of worthwhile weblogs is substantially higher than the 5% Sturgeon’s Law might suggest.

Here’s a list of 25 people: Steven Bell, David Bigwood, Laura Blalock, Susan Crawford, Walt Crawford, Anna Creech, Lorcan Dempsey, Bill Drew, John Dupuis, Elizabeth Edwards, Edward Felten, Geoff Harder, Sarah Houghton, David King, Elizabeth Lane Lawley, Carole Leita, Lawrence Lessig, Andrea Mercado, Mary Minow, Christina Pikas, Aaron Schmidt, Kristina Spurgin, Peter Suber, Sheila Webber, Donna Wentworth.

The list is presented alphabetically for lack of a better organizing principle. Women make up almost exactly half of the list. That wasn't intentional, but it's clear that women are better represented among thoughtful library bloggers (and thoughtful copyright-related bloggers and thoughtful lawyer/bloggers and...) than they are in some more traditional areas of technology visibility.

I'd say 19.5 of the people on this list are primarily "library people," 5.5 aren't. (Geez, Liz, I don't know: What do you consider yourself these days?) A few of the people on the list are older than I am. Several are young enough to be my children.

Roughly half of the people on the list have published books, at least one of them more than I have. Most of the people on the list have published articles—scholarly, popular, or both. Those that haven't almost certainly will, because that's one of several things these 25 people have in common: They're articulate, thoughtful, professional, and write things that are worth reading—at least some of the time. (Hey, none of us bats 1000.) Those who haven't yet had substantial impact on their fields almost certainly will—there's not a name on the list from whom I wouldn't expect significant accomplishments. (Some already have admirable track records, to be sure.)

Do any of these people call themselves "blog people"? Some do, some don't. But they all blog. Every one of them.

I could easily list another couple of dozen, including many of the names mentioned earlier in this essay. (Sorry, George, Dorothea, Blake, Rochelle, Michael, Simon, Steven, Jenny, et al. No insult intended: I wanted to keep the list reasonably short.)

Michael Gorman is no fool, no Luddite, and no idiot. On the topic of blogs and their writers, Michael Gorman is also no authority.

Trends & Quick Takes

Critiquing the Curriculum

That's the title of Wayne A. Wiegand's article in the January 2005 *American Libraries* (pp. 58-61). The sentence beneath: "The entrenched LIS agenda needs to change to reflect the most critical functions of the library." This is a case where I need to remind you that IANAL—in this case, L for Librarian: I haven't been to

library school. And maybe, if Wiegand's right, that explains why I wrote *Being Analog*, cowrote *Future Libraries*, and gave some of the speeches I did: I wasn't aware that there was an entrenched LIS agenda.

Wiegand's first point is the one that affects me most directly:

First, I'm convinced that most [library school faculty] think of libraries as part of a greater world of information... However, my study of American library history leads me to see information as only part of a larger library world, in which libraries have done three things especially well for the past century-and-a-half: They have 1) made information accessible to millions of people on many subjects; 2) provided tens of thousands of places where patrons have been able to meet formally as clubs or groups, or informally as citizens and students utilizing a civic institution and a cultural agency; and 3) furnished billions of reading materials to millions of patrons.

The library as place. *Reading*—and reading materials in general, not just "information." As Wiegand recounts, in library-related questions on a 2001 national survey, when people were asked what skills librarians most needed, 76% said "familiarity with a range of books and subjects"; when asked what people do at public libraries, 92% said "borrow books."

In some of my talks, I've dissected one ALA slogan, "The information place," as being wrong on all three counts. First, libraries have never been *the* place people get information—and mostly not the primary place people get either up-to-the-minute information or the information most important to their careers and hobbies (at least for public libraries). Public libraries fill in the pieces: They provide access to the information people don't acquire as a matter of everyday life. Second, "information" isn't all that libraries are about—I don't even believe it's *primarily* what libraries are about (and Wiegand seems to agree). Third, while the library as place is vitally important, libraries have served beyond their walls for many years. Wiegand says library schools need to understand more about place and reading, and why they're so important to patrons. I think that's true.

The other major point is that LIS faculty are inclined to "think primarily of...the user in the life of the library" and that they need to think of the library in the life of the user. That's certainly true, but it's an area for others to comment and expand on.

For all I know, Wiegand could be wrong: Maybe library school faculty have already dropped infocen-

tric attitudes and started to understand the importance of reading, place, and the library within the overall life of the user. I suspect he's right. **Go read the article** (if you haven't); it's worth thinking about.

A Safer Internet Explorer Will Destroy Firefox?

The February 15 story at News.com is straightforward enough: Internet Explorer 7 will appear in beta this summer, rather than waiting for the next major version of Windows (as would normally happen). Microsoft appears to recognize that IE represents one of several major vulnerabilities in Windows.

Naturally, analysts assumed that Firefox has something to do with Microsoft's change of plans: For the first time in many years, thanks mostly to Firefox, IE has less than 90% of the PC market.

So far, so good—but Molly Wood, senior editor at CNET.com (News.com's parent) extrapolates in a manner I find disturbing and unsupportable. The article title almost says it all: "IE 7: so much for Firefox." The first paragraph: "The party's over." Why? Because "Papa Bill just dropped the hammer." Here's Wood's take on Firefox:

Firefox is great. I use it. But it's a chore sometimes, what with most sites using that pesky nonstandard IE code. Not everything renders properly, and some sites just plain don't work—I have to load up IE to use them. Plus, let's be honest—Firefox has its flaws. Why is there no way to check for updates from within the browser, for one thing? Why does it take so doggone long to launch? Why, *why* must it crash every single time I open a PDF? I mean, every *single* time...

I wonder about two things: That first sentence (belied by everything else in the paragraph) and just what's wrong with Molly Wood's PC. I'm running Firefox on a middle-aged PC, dialup, at home (sometimes) and on a recent underconfigured PC at work (all the time except for one solitary website requiring IE). Rendering improperly? I have yet to see it. Sites that don't work? Yes, some sites are coded so they'll reject any browser except IE—but not many. No way to check for updates? Just the other morning, Firefox alerted me via popup window and icon that there were significant updates: no *need* to "check" for them. "Take so doggone long to launch"? One reason I use Firefox most of the time is that it's *faster* to launch (and use) than IE. Crashing on PDF? Well, it's never happened to me, and I use PDF a lot. What am I doing wrong?

Given Wood's ability to turn "great" Firefox into a series of disasters, I'm not surprised that she concludes, "If a standalone IE 7 is even 50 percent more secure than current versions, the Firefox rebellion is finished." Oh, and if IE 7 has tabs, Firefox "will be destroyed as surely as the Hungarian uprising of 1956 was crushed by Russia."

Hmm. Maybe there's a clue in that last statement. Remember what happened in Hungary in the long term, admittedly after 1956?

iPod Ubiquity

On one hand, here's another Pew Internet report—asking how many American adults have iPods or MP3 players. The answer, if you assume that 2,201 people who are willing to respond to phone surveys represent a valid statistical cross-section: 11%, more than 22 million of them. 14% of men, 9% of women; 19% of those under age 30 (noting that with each subdivision, statistical validity weakens), 14% of "younger Baby boomers" (40-48). "Upscale" households (\$75,000 or more household income—hardly upscale in these parts, but never mind show 24% penetration, as compared to 10% of \$30K-\$75K households and 6% of under-\$30K households. If you have broadband at home there's one chance in four that you'll have an iPod or MP3 player—but only 9% of those with dialup connections do. (What about those without internet connections?)

On the other, there's a March 28, 2005 *Media Life* piece, "Not everyone owns an iPod." Lorraine Sanders notes, "[Y]ou'd think half the country's teenagers own an iPod, that TiVo is taking down primetime television, and that Howard Stern's impending move to Sirius will spell the end of traditional radio as we know it." She offers a paragraph to comment on that hype:

Most of this is simply hogwash.

According to an Arbitron/Edison Media Research study, more than 80% of Americans plan to continue listening to on-air radio; 6% own or use a TiVo or other DVR; 3% own a Blackberry-equivalent—and only 6% own or use an iPod. Consumers with the ability to block ads don't block all of them—and that's only reasonable.

Note that both studies *could* be right—if MP3 players other than iPods represent around 45% of portable-player sales, which is quite possible if you include flash players as well as disc players.

PublishAmerica: Vanity or Not?

An Associated Press story on January 29, 2005 talks about PublishAmerica—founded by a “Web marketing consultant who had written two books he couldn’t get published” and a client who owned a vanity press. The new venture “would take on those people who yearned to be authors but struggled to find a publisher, offering the editing and promotional support not found at a vanity press and do it without a fee.”

And boy, have they done it: more than 4,000 books released in 2004, with nearly 11,000 writers “under contract.” The website says that signing with PublishAmerica gives authors “the very important distinction of having your next book accepted by a traditional publishing company” and assures applicants that manuscripts are carefully reviewed and edited, that books are available in stores, and that authors don’t have to pay. And, to be sure, founder Larry Clopper (the unpublishable Web marketer) says, “The publishing industry will never be the same.”

One novelist who couldn’t find a publisher loved being accepted by “traditional publisher” PublishAmerica—but says her manuscript wasn’t edited, there was minimal marketing assistance, and bookstores told her they don’t stock PublishAmerica books because they don’t consider it a real publisher. One watchdog group calls the company an “author mill.”

Consider these numbers: In 2004, PublishAmerica released 4,800 books. Gross revenues totaled “\$4 million to \$6 million”—in other words, no more than \$1,250 per book, an amount barely sufficient to cover good editing and manuscript preparation, much less marketing and fulfillment. And that’s gross: PublishAmerica claims to have 70 full-time staff, which at (say) \$30,000 average salary would mean \$2.1 million of that \$4 to \$6 million is pure overhead. Assuming an average price of \$20 per book, PublishAmerica may have sold as many as 300,000 books—or about 63 per title. The “all time best seller” has sold around 5,200 copies total.

Author advances range “from \$1 to \$1,000”—and since there’s no up-front fee, PublishAmerica isn’t a traditional vanity publisher. Or is it? “[B]ecause PublishAmerica has little clout in the market, authors end up buying copies from the publisher, which periodically offers special discounts, and selling the books themselves.” The best the head of the company will claim is that sales to authors are “less than 50 percent.” And, despite his claims of PublishAmerica’s

“traditional” status and the website’s claims of availability in stores, he can’t cite *any* PublishAmerica books that have been placed in bookstores nationwide. Instead, almost all PublishAmerica production is print-on-demand, where a book doesn’t exist until it’s ordered.

Then there’s the careful editorial review. One author was unhappy with the handling of a novel, so she submitted a new manuscript: the first 50 pages of the previous novel followed by the last 10 pages, repeated often enough to make a manuscript. The manuscript was accepted. Another writer submitted a novel consisting of 30 pages repeated six times: it was accepted.

The Science Fiction and Fantasy Writers of America won’t offer membership to PublishAmerica authors; neither will the Authors Guild or the Mystery Writers of America. PublishAmerica had its own comments about science fiction authors (and, full admission, *right here it gets personal*, even though I learned more than 40 years ago that I have no talent for writing science fiction):

As a rule of thumb, the quality bar for sci-fi and fantasy is a lot lower than for all other fiction... [Science fiction authors] have no clue about what it is to write real-life stories, and how to find them a home... [They are] writers who erroneously believe that SciFi, because it is set in a distant future, does not require believable storylines, or that Fantasy, because it is set in conditions that have never existed, does not need believable every-day characters.

Wow. Take that, Gene Wolfe, JRR Tolkien, Terry Pratchett, and a few thousand others: You can’t compare to “all other fiction” (which presumably includes romance, mystery, and porn).

Thirty science fiction writers found this a bit outrageous, so decided to see just how tough PublishAmerica’s standards really were. They got together over a holiday weekend and wrote *Atlanta Nights*, “a novel about hot times in Atlanta high society,” making it as bad as they possibly could. As you’d expect, PublishAmerica accepted it—and, of course, withdrew the offer “upon further review” after the writers said what they’d done. (You can find *Atlanta Nights* at www.lulu.com/travis-tea.) [Later information from an SFWA January 28 press release on PRWeb.]

Quicker Takes and Mini-Perspectives

I was a bit surprised by a mini-roundup of “gaming notebooks” in the February 8, 2005 *PC Magazine*. Not by the prices—hey, if you’re a gamer, what’s \$5,300 or

\$5,600? Not by the brandnames—Alienware, Falcon Northwest, and Voodoo are all familiar names for this type of PC. What surprised me is what's not mentioned consistently. None of the reviews mentions the size of the computer, and only two of three even mention weight (which in both cases is a tad heavy for a portable: 8.8lb. for the Alienware, 12.4lb. for the Voodoo). Don't weight and size matter when you're paying a hefty premium for portability?

- Arggh. I can't take it any more. Here's an Editors' Choice review of the Logitech Z-2300 "2.1" (stereo with a subwoofer) speaker system, \$150, in the March 2005 *Computer Shopper*. The first two sentences of the third paragraph sent me over the edge: "Our first test on the Z-2300 was musical, and we loved the results. The system delivered the kind of thump and low-end resonance you'd associate with a great PA system at a concert." As opposed to a system that accurately delivers the bass that was in the original recording, which probably isn't "thump and low-end resonance." Other comments about this speaker, lauded for its "outstanding audio quality," are that game play "made our walls, desk, and floor shake with an intensity our neighbors quickly hated" and that, viewing a *Lord of the Rings: The Two Towers* DVD, "every collision benefited from a tremendous thud, and the fall of every horse hoof came through with pristine clarity and low-end bass that made us shudder." Good grief. Flatness of response? Harmonic distortion? Accuracy? Never mind: It shakes those walls, and that's what matters.
- The print book continues to die slowly if at all. A February 24, 2005 AAP press release says that net sales in 2004 totaled \$23.7 billion, a 1.3% increase from 2003—not a big increase, but not bad for a mature industry in a society where "nobody reads books." Adult hardbound sales increased a respectable 6.3%, with paperbounds up 2.8%. Compound growth rate from 1997 through 2004 averaged 4.7% per year.
- If the classic DVDs you buy from Warner Brothers seem unusually good, there's a reason, according to Fred Kaplan in a February 28, 2005 *Slate* story. For one thing, the film libraries owned by WB (which include RKO

and pre-1986 MGM films) have been well preserved, and most DVDs are mastered from the original negatives. For another, Warner has restored its true Technicolor movies (the ones filmed on three black-and-white strips between 1935 and 1954), going back to the triple b&w negatives and aligning the negatives properly. Since b&w negatives don't fade (unlike color), this combination of archiving and high-resolution digital scanning and frame-matching can yield spectacular results. Additionally, WB tends to produce most classic movies as two-disc sets, putting the extras on the second disc—which means they can use milder compression rates on the movie disc, yielding better pictures. So if *Meet Me in St. Louis*, *Gone with the Wind* (where the movie's spread over two discs), and the forthcoming release of *Wizard of Oz* look better than you remember, it's because they probably are better than you remember.

- John N. Berry III of *LJ* wrote about some flack the journal received last year after quoting posts from lists. "Before publication, we decided it was unnecessary to get permission to quote these posts, since a large audience had already received their messages." That's an interesting definition of "large"—some 1,500 subscribers for one list, but only 400 or so for the other. But that's not the point. This is: There was an "e-storm of messages" on one list saying the quotes represented "bad journalism, bad netiquette, and bad ethics to quote people without even telling them, or asking their permission." I know it's been a long delay—but I'm with John on this one: "[T]hose who go public with their thoughts...have some responsibility, too." If you sign your name to a list post, especially if it's a list anyone can subscribe to, you've *published* that post—and it's neither unethical nor bad etiquette to excerpt from a publication without asking permission. (If an entire post was quoted, that's a different issue: the poster could conceivably sue *LJ* for copyright infringement, since every signed post is automatically copyright-protected. But if you haven't registered the post, you can only sue for actual damages, and those would be

mighty tough to prove.) I find very few bloggers who won't quote from list postings, and many bloggers want to claim that they're providing a form of journalism. I don't believe you can have it both ways. I was surprised by a later letter in *LJ*, where the writer asserted that there's a distinct difference between quoting published material and using material from a list, claimed that lists go to "a defined group of people who share a common interest" and called selective quotation "steal[ing] their words." Nope, sorry: list postings are published material...unless the list is *factually* private and limited to those who can prove a common interest.

- The proprietress of *Exploded Library* offered some "possibly luddite confessions" on February 3, 2005. I was charmed. She's never bought or sold anything on eBay; she likes to subscribe to the print version of her urban daily newspaper; and she prefers to use travel agents to book flights—many of which are complicated international trips. Well, you'll have to call me 67% Luddite as well, then—and maybe 100% for *complicated* flights. I've spent even less time on eBay than she has. I regard a good metro daily newspaper as important to staying socialized, staying aware of all those things that aren't crucial to my daily life but might matter in the future. I book most of my own flights, but most of them are simple. We do use a travel agent for cruises—and would for any complex trip.
- David McCandless wrote an interesting confessional in dot.life at *BBC News* on February 7, 2005: "Why I'm giving up broadband." He's a "committed early adopter and geek," has used broadband for years—and has come to realize that "there isn't really that much I can do with broadband." His wife convinced him not to play online games too often; he's noted that you can get full-motion video of news just as easily by turning on the TV; and he doesn't want to discuss illegal file-sharing. He believes he's formed "an information habit" that's interfering with the rest of his life. So he dropped back to dial-up. Which works just fine for e-mail and web browsing. "Isn't that what the internet is really for?"

- I haven't spent much time exploring it, but you might check out *Low Threshold Applications* (jade.mcli.dist.maricopa.edu/lta/), a site for applications of technology (largely for educational purposes, including library-related purposes) that are "reliable, accessible, easy to learn, non-intimidating and (incrementally) inexpensive." As of early February, there were 44 such applications on the site, from "integrating RSS feeds into your course management system" to "creating editable forms in Microsoft Word."

The Good Stuff

Flanders, Vincent, "The biggest web design mistakes of 2004," *Web pages that suck*. www.webpagesthatsuck.com

Vincent Flanders runs the site noted above with "Daily sucker" and other features. As guides to really bad web sites go, WPTS is less snarky than some—and is itself, as Flanders cheerfully admits, an example of some sucky techniques. This 15-page essay (posted in two parts) results from his examination of a year's worth of Daily Suckers.

It's **well worth reading** to consider what you might be doing dreadfully wrong. His mistakes (each discussed in considerable detail, with links to examples if you're reading online):

- Believing people care about you and your web site.
- A man from Mars can't figure out what your web site is about in less than 4 seconds.
- Mystical belief in the power of Web Standards, Usability, and tableless CSS.
- Using design elements that get in the way of your visitors.
- Navigational failure.
- Using Mystery Meat Navigation.
- Thinking your web site is your marketing strategy.
- Site lacks Heroin Content.
- Forgetting the purpose of text.
- Too much material on one page.
- Confusing web design with a magic trick.
- Misusing Flash.
- Misunderstanding graphics.
- AFFront page.

A few more words about some of those. He's not saying anything's wrong with standards and usability—only that following them does not guarantee that your site will be any good. Mystery Meat Navigation, one of Flanders' pet peeves (unless it's used on music, art, movie or fashion sites and the like), is when you need to mouse over unlabeled buttons or other graphics to actually find pages in the site, as JavaScripted text shows up. Heroin content is content that your audience *really wants badly*. "Forgetting the purpose of text" deals with a peeve that's one of my pets, sites where designers have made content secondary to appearance, at the expense of easy reading (or printing!). And an AFFront page is a home page that affronts you—and "odds are it was created by Microsoft's (Af)FrontPage." Flanders notes that Microsoft doesn't use FrontPage for its own site—not even on the pages describing FrontPage. That's only a bit of what's in this two-part essay.

Nellis, Michael, "What constitutes 'information'?" *The lair of fang-face dreamweaver*, August 10, 2004. www.angelfire.com/scifi/dreamweaver/bannedbks/

Michael Nellis comments actively at LISNews (under his "fang-face" persona) and uses an exchange on that site as the basis for this 19-page commentary. "Just what is information and how can you tell if it is valid or not?" Another LISNews pseudonymous correspondent (Tomeboy) wrote that Alternet.org "was not a valid source for information because it was not listed at [LexisNexis]," to which Nellis replied that some Alternet articles were reprints from mainstream newspapers "and that [Tomeboy's] problem with Alternet.org was simply that most of the articles there did not support his prejudices."

But the point of the discussion is the opening line. Nellis talks about "information" as "an accurate reflection of reality"—then goes on to look at belief systems, facts, disinformation, misinformation, and opinion. People get into trouble when they regard their opinions as facts and when they add disinformation to the mix.

There's a lot more here. Nellis offers some useful commentary and quotations on how we test ideas for ourselves. It's mixed in with some strongly stated political opinions, and this is not a tightly edited or spell-checked essay, but you might find it worth reading. (If you're a big George Bush supporter, you probably won't make it through the first six pages.)

Richards, Linda L., "Just because you can...", *January Magazine* January 2005. www.january-magazine.com/nonfiction/selfpub.html

This review essay covers four books on self-publishing—or, rather, two on self-publishing and two on publicizing books. More than that, Richards—editor of *January Magazine* (which reviews a lot of books)—discusses the whole issue suggested by the essay's title. *Should you self-publish?* Self-publishing isn't the same as using a vanity press. With self-publishing, you know you're the only one who's going to publicize your book or get it into bookstores (if that's even possible), you're the one who's going to try to get the book reviewed—and you're the one who will invest all the front money and realize whatever profits ensue.

As Richards points out, self-publishing used to be a "big hairy deal": hiring a contract editor, getting a design house to design and typeset the book, contracting with a printer for the press run, which in the days of offset lithography probably meant opting for "at least a few thousand books" because the first-copy cost of typesetting, platemaking, and setup was so high. Then, with a garage full of books: "What, you'd have asked, do I do now?"

Things have changed. You can do the whole process in your home office, which would "take some pretty specialized gear" but might require less investment than contract editing, much less the whole typesetting/lithography process. More probably, these days, you'll prepare the book's content, typography, and cover design on your home computer, then "simply [send] your prepared files to a print on demand publisher who will print as many or as few of your books as you want for relatively little cash." Or even have them handle online sales for a bigger cut of the revenue—leaving you with *no* books in hand other than the ones you want for yourself and your friends.

January Magazine reviews some self-published books and Richards loves to see the good ones, but she notes that most of them are "not so good," with some "downright bad." Not just bad concepts and storytelling, but "books so poorly conceived and executed you have a hard time getting to the meat of the topic at hand." She believes that part of the problem is that it's so easy to make a book—but notes that there's no reason to produce a bad book. Then she goes on to review the books at hand. One of them, oddly, doesn't follow the advice within the book (and it's bad advice

anyway: You really shouldn't deface review copies). Still, the book—*Creative Self-Publishing for the World Marketplace* (Marshall Chamberlain, Grace Publishing, 2004) “includes small amounts of advice across a truly wide and potentially bewildering array of topics” and is “a reference that every self-publisher should have on their desk.”

If you're interested in self-publishing or PoD publishing (essentially the same thing, but traditional publishers also use PoD in some cases), **read this essay**—and you should probably be reading *January Magazine* as well.

Silverthorne, Sean, “The hidden cost of buying information,” Harvard Business School *Working Knowledge*, November 8, 2004. hbswk.hbs.edu

We all need good information to make decisions—that is why consulting is an industry that never goes out of style. But paying for information can carry a hidden cost: We may give it more weight in our decision making than it deserves.

That's the lead paragraph in an interview with Francesca Gino, a Harvard Business School post-doc who recently published a working paper on “overweighing” information because it's expensive. She did an experiment where subjects were asked to answer different sets of questions about American history. Subjects had the opportunity to receive free advice or expensive advice—and they were *told explicitly* that the quality of information would be the same in all cases. (All advice came from the same source.) She found people tended to pay significantly more attention to expensive advice than to free advice, even if they had reason to believe it was equivalent.

I'd like to say I'm surprised, but I'm not. The full interview offers more background and is **worth reading** (it's only six pages). Among other things, this phenomenon would explain why businesses hire consultants to tell them what they already know from internal analysis: Paying for the information makes it more valuable. Gino talked to a dietician who noticed that patients were more inclined to follow a diet obtained at high cost than a diet (even the same diet) available for free.

Maybe it's partly because we're “over-informed”—or at least faced with too much possible information. Maybe it's because “you get what you pay for” is so ingrained in our culture. Payment doesn't have to be money, of course: An investment of time or effort also makes information seem to be more valuable.

I've wondered whether this is a problem for open access publishing—that free journals will tend to be regarded as less worthy than overpriced journals. I'm almost certain it's a problem for the valuable free gray literature of the web, such as this here publication.

Slattery, Oliver, Richang Lu, Jian Zheng, Fred Byers and Xiao Tang, “Stability comparison of recordable optical discs—a study of error rates in harsh conditions,” *Journal of Research of the National Institute of Standards and Technology* 109:5 (September-October 2004): 517-524. www.nist.gov/jres

The title's almost longer than this brief mention. If you're interested in medium-term survival of CD-Rs and DVD-Rs, **read this article**. You won't get brand recommendations, but it's one of the toughest testing programs I've heard of, exposing recorded discs to accelerated aging tests and studying the error rate over time. Results show that phthalocyanine dye (which some Verbatim DataLifePlus CD-Rs use) performed better than other dye types, that this dye combined with gold-silver alloy as a reflective layer consistently offered the best stability and could yield semi-archival data (several tens of years, not centuries)—but that it's difficult to identify stable media and that direct exposure to sunlight on the rainbow side may be the fastest way to render a CD-R or DVD-R useless.

Masthead

Cites & Insights: Crawford at Large, Volume 5, Number 6, Whole Issue 62, ISSN 1534-0937, a journal of libraries, policy, technology and media, is written and produced by Walt Crawford, a senior analyst at RLG.



Cites & Insights is sponsored by YBP Library Services, <http://www.ybp.com>.

Hosting provided by Boise State University Libraries.

Opinions herein may not represent those of RLG, YBP Library Services, or Boise State University Libraries.

Comments should be sent to wcc@notes.rlg.org. *Cites & Insights: Crawford at*

Large is copyright © 2005 by Walt Crawford: Some rights reserved.

All original material in this work is licensed under the Creative Commons Attribution-NonCommercial License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc/1.0> or send a letter to Creative Commons, 559 Nathan Abbott Way, Stanford, California 94305, USA.

URL: cites.boisestate.edu/civ5i6.pdf