Cites & Insights

Crawford at Large

Volume 4, Number 11: September 2004 ISSN 1534-0937 Walt Crawford

Bibs & Blather

Finding It

Sometimes online items cited in *Cites & Insights*—either directly, with a citation, or indirectly within an article—have URLs (either full ones or ones leading to the general site). More frequently these days, they don't. One might argue that this is unforgivable sloppiness. One might be right, but the case could also be made that URLs change, that some of these items come from database-drive sites—and that the information provided in a citation should almost always be enough to locate the article.

As the sloppy author/editor/publisher here, I favor the latter argument. I don't always have the URL for a paper when I'm working with it; PDFs only include a URL if the author chooses to provide it within the body of the paper. Quite a few URLs are simply too unwieldy to include—you can't act on them directly anyway—and I'm not inclined to add URLs in an article except in unusual circumstances.

Consider the first item under "Brief Commentaries" in the LIBRARY ACCESS TO SCHOLARSHIP section in this issue. I have a three-page printout that begins "After the Tipping Point – What Next?" and ends "Adam Hodgkin, President and Co-Founder xrefer, [a description of xrefer], "14 May 2004." No URL. But the search "hodgkin 'tipping point'" in Yahoo and AllTheWeb pops up an archived copy of the article as the first result; while Google doesn't do as well, its first result is an Open Access News weblog entry that links to the article. So while I should perhaps apologize for the slovenly nature of web references in *Cites & Insights*, it's a limited apology. There may be such a thing as a free lunch, but the service may not be up to snuff.

Walt Crawford with a Weblog?

Not really—but I apparently made Steven Cohen's day anyway. I *did* start a Blogger weblog at cical.blogspot.com, but it's "C&I Updates," pointing to cites.boisestate.edu as my homepage (I may add a link to *my* homepage at some point), and its real purpose is to provide an RSS feed for announcements of new *Cites & Insights* issues (and, once in a while, related announcements such as expansion or modification of the overall direction, or developments in monetization). The full-text Atom feed (Atom, RSS, I don't understand the difference, and Bloglines for one doesn't much seem to care) is http://cical.blogspot.com/atom.xml, if you don't even want to visit the vanilla blog site once.

I can't imagine anyone adding C&I Updates to their bookmarks/favorites; there most assuredly won't be daily postings, words of wisdom, or hot controversy. No blogroll. No fancy set of links. It's there for the Atom feed (and figuring out how to add that Atom feed took a lot longer than setting up the weblog); it's *intended* for people who would rather receive issue announcements via aggregator than via email. Given the state of email these days, I can't argue with that position. (If enough people migrate off Topica to the feed, I'll shut down the Topica announcement list, but I'm in no hurry to do so.) By the time you read this, a link to the Atom feed (and to the weblog) will also be on the *Cites & Insights* home page.

Inside This Issue

I do have an odd little journal at LISNews, but it's a journal, not a weblog, and it's mostly for nonsense that doesn't belong in *C&I* itself.

Speaking of Being Offtopic

I wasn't, but a good non sequitur seemed in order at this point. Space allowing, I'll do another "Offtopic Perspective" in the next issue, commenting on the *rest* of the movies in the *InsideDVD* 40-movie starter pack—well, all but one where after 15 minutes I was unwilling to continue. I've got four left (and 33 minutes of a fifth), two of them short anti-marijuana "public service" movies (that's right, *Reefer Madness* is on the agenda—in glorious black and white, not some colorized version).

The original Offtopic Perspective talked about making do with (for example) the many commentary tracks on *Lord of the Rings*, and the like, after I run out of the 90 movies on hand at that point. While I do interrupt the flow of old movies for *Buffy* commentary tracks as we encounter them, "running out" doesn't seem like a threat in the near future. Thanks to Seth Finkelstein, I found that Overstock.com has a range of other 50-movie Megapacks: 12 two-sided double-layer DVDs, each with three or four hours of stuff on each side, in a compact box, typically selling for $28. I've already mentioned the "Family Classics" megapack that I picked up for $20. I think there are at least six more now—and they're pretty clearly still being produced. Some have 100 TV episodes or 150 movie serial episodes instead of 50 "movies" (some movies are collections of shorts). I picked up the "sci-fi" megapack (some of these gems must surely have been on *MST3K*!) and, just this week, an "all-star" megapack in which every movie is in color. There are also western, horror, mystery, and comedy megapacks (but the comedy one didn't entice me).

I wonder whether some of these might make interesting low-cost additions to public library DVD collections, as very cheap ways to provide collections of miscellaneous older movies (although they're not all that old): Buy two copies, circulate the 12 cardboard disk sleeves separately (four or five movies on each, with blurbs on the sleeves), and toss 'em when they wear out or disappear. (Or make copies, for that matter—most or all of these should be in the public domain, although *this is not legal advice*. But then, if sets are still available at $28 for 12 discs, why would you consider burning four $1 DVD-Rs to replace a $2.35 DVD?)

So I now have 150 movies to go after the current pack is complete. That's two years of exercising, and I suspect the company will produce even more cheapo collections. There must be other sources for the Megapacks (which still seem to have a $199 "original list price"), but Overstock's where I've found them.

Perspective

The Quality Contradiction

Super Audio CD (SACD) and DVD-Audio both offer sound quality that may be significantly better than CD Audio. Both are struggling to make headway in the market, while sales of devices to play low bit rate MP3 files—audibly degraded from CD quality to anyone with halfway-decent ears—take off like crazy. I'd guess that legal downloads of degraded-quality MP3 and AAC music outsell SACD and DVD-A, even leaving out illegal downloads.

The general quality of regular TV sets has improved substantially, DVD movies offer roughly twice the picture quality of VHS, and high-definition TV offers much larger improvements in picture quality. Meanwhile, the storage capacities of most personal video recorders are advertised based on a speed at which recording quality isn't as good as VHS; movie studios worry about losses from illegal downloads of movies with picture quality *considerably* inferior to VHS; and the adoption of S-VHS recording (which offers 60% better picture quality than VHS) near the end of VHS's reign as a primary medium was roughly the same as it was two years after S-VHS was introduced: two to four percent.

What's going on here?

Quality and Convenience

Life is rarely an either-or proposition, and that might be a good enough answer. I think there are at least three other aspects of this apparent contradiction, two of which I'll discuss here. (The third, misleading sales and advertising, is truly out of scope.) The first is that people sometimes choose convenience over quality, at least in some areas. That's both natural and sensible, although it helps to be aware that you're making the choice. If you want music while you're jogging, a solid-state MP3 player or even one based on a hard disk (like the iPod), but only using the hard disk once every few minutes, is likely to work a lot better than a CD player—and if it's a solid-state MP3 device, storing the songs at a sub-FM-quality 64kb rate is tempting, since you get twice as much music as at FM-quality 128kb.

I believe that's how audiocassettes came to challenge and eventually surpass vinyl sales. They didn't sound as good (if you took care of your vinyl), but they were a *lot* easier to handle and made your music portable. "If you took care of your vinyl" was a significant hurdle for a fair number of people. The process required to keep records and expensive styli in good condition can be daunting. (Remove the record from the inner sleeve—having in some cases replaced the inner sleeve with a better quality sleeve—without touching the grooves. Clean the record with an appropriate special liquid and either a handheld brush or, better yet, a vacuum cleaning system. Then use a static gun to eliminate stored static. Then clean the stylus with a special brush and liquid. Then you can play the record! I did this faithfully; my 1,200-odd records were generally in like-new condition when I got rid of them.

CDs combined quality and convenience, although the debate over whether CDs sound as good

as the best vinyl may never be settled. (Many early CDs sounded *awful* because they were remastered from master tapes that had been "mixed hot" to sound lively on cheap record players; the resulting high end was painful to the ears.) For us, it was easy: The first new CD we purchased (*Graceland* by Paul Simon) sounded great, as did the first replacement for an LP, and there was no hassle: Take CD out of case, put in player, push button. We never purchased another LP, and I've never regretted that decision. (If you believe vinyl is better, more power to you. You'll be pleased to know that sales of vinyl LPs *and* turntables have been increasing for the past few years, although most turntables and cartridges are incredibly expensive. But then, if you really believe in vinyl, you already know all this.)

Most of today's CDs sound better on most of today's CD players, even $20 portables, than most early CDs sounded on most early players. Are they "perfect sound forever"? No, and I don't believe Sony itself believed that early slogan. For some of us—myself included, I suspect—good CD sound represents all we'll ever be able to hear in two-channel sound. But some people do hear differences between stereo CD and stereo SACD or DVD-A, even in the double-blind tests that high-end stereo writers disdain. I'm satisfied that it's a real difference for some people. The other selling point of DVD-A and, in some cases, SACD is multichannel sound: A failure back when three semi-compatible systems were battling it out for "quadraphonic" honors and most people didn't want all those speakers in their living room, but more possible now that millions of people *already* have surround speakers installed for DVD movies and HD television.

You or your library may own some SACD discs already, without being aware of it. Sony and a couple of other companies are being clever on some major re-releases such as (some) Rolling Stones and Bob Dylan albums: They're coming out as dual-layer SACD discs, with one layer offering higher-resolution and possibly multichannel sound, the other layer a standard CD that plays in a standard CD player. If the SACD logo is on the disc at all, it's not emphasized, and there's no pure-CD equivalent. (If you ignore those guerilla-SACD discs, vinyl LP sales actually total more than SACD and DVD-A combined, according to some reports.)

I'm not saying DVD-A and SACD have failed. They haven't and probably won't, but they also haven't succeeded nearly as rapidly as analysts projected or companies hoped. Recording companies would *love* to switch to high-resolution formats, not only because they might be able to sell you the same music one more time, but also because SACD and DVD-A both come with built-in copy protection, unlike CD audio. Dual-layer SACD/CD hybrids are not, typically, copy protected on the CD layer.

But they sure aren't succeeding as rapidly as MP3 formats—and low-resolution MP3 *absolutely* loses some of the sound quality of CD unless you select at least 196k and possibly 320k or higher data rates. When you're ripping CDs to MP3 for use in portable devices, that's fine: You're making a choice for convenience and can always go back to the original. When you're paying for legal downloads, it's not so great: Expanding 128K MP3 to CD audio form does not restore the lost sound quality. Once it's gone, it stays gone. The convenience choice precludes a later preference for quality—if you think you'll ever care.

I'd like to believe that DVD's rapid ascent has a lot to do with picture quality—but if that's true, then why was S-VHS such a dud? I suspect that the sheer convenience of DVD has more to do with it—that, and the extras that come with DVD. (Some people care a lot about the restoration of the original picture in widescreen DVDs, but I think those are mostly the same people who care about the picture quality: A substantial percentage of us, but certainly not everybody.) If that's true, then high-definition DVD (either of the two competing formats likely to emerge this fall or next year) may be in trouble: It won't be more convenient, and people may not care about the quality.

If you have a "40-hour" TiVo, do you record at a rate that puts 40 hours on the disk or at a 20-hour or 10-hour rate? It's more convenient to be able to store more; do you care about the degradation of picture quality? If you burn TV or home movies to DVD+R/-R, do you put one hour on a disc, two, or four? Are you trading convenience for quality?

Noticing the Difference—and Caring

Maybe you answered "40 hours" and "four hours" to the first and third questions in the previous paragraphs and don't understand the second and fourth questions. "What degradation? The picture looks fine to me." Similarly, you may think there's really no difference between "FM quality" and "CD quality," and that 128K MP3 *is* CD quality, with 64K "good enough."

There are two related issues here: Noticing the difference and caring about the difference.

Some people just don't notice differences in some areas. I think we're all more sensitive in some areas than in others. Women generally hear better than men do (and tend to be particularly sensitive to some forms of distortion). Old farts like me usually have degraded high frequency hearing. I don't claim to be a discriminating judge of fine food; my taste

buds aren't that sensitive. I don't know that I'm really a connoisseur of fine wine, although I can certainly appreciate some differences. I'm *certainly* no connoisseur of perfume (and generally avoid it as much as possible). Although I believe I understand differences in car performance and appeal, we own and drive Honda Civic EXs by preference. I probably don't appreciate the differences in high-end clothing. There are loads of areas in which I don't pay attention to, or even understand, the differences between the good and the best.

We *do* see the difference between good broadcast/cable TV and the same show recorded on regular VHS, even at full speed; that's why we've never owned anything but an S-VHS VCR or used anything but S-VHS to tape shows. I've never understood why more people didn't see the difference. But maybe that's the wrong issue. Maybe they *see* the difference but don't care. Or at least maybe that's been true for 96% of VCR purchasers. Even at a price increment of $50 or so, S-VHS never made inroads in the marketplace.

Even with my mediocre hearing and inexpensive speakers and headphones, I can tell the difference between a typical CD and 128K MP3. Apparently, most technology writers and other journalists can't and neither can many other people. Most of the audio CDs I listen to are mixes based on MP3 storage—but the MP3 files are ripped at either 196K or 320K, using current Frauenhofer codecs in MusicMatch Plus. I'm not *sure* whether I can tell a difference between 196K and 320K, at least not sure enough to re-rip all the old CDs, but I'd certainly rip any classical music at 320K and I rip any new CDs at that rate. My wife and I both hear enough audio differences so that, when we decided to buy a modest music system to replace the old speakers that were too big for our little house, we chose a $700 system over some impressive $300 and $400 systems. The $700 system didn't get in the way of the music and seemed to reproduce significant differences among tracks; we had little trouble agreeing that it was worth the extra money.

You might not make those distinctions—or you might not care. The same goes for TiVo and other PVRs: Maybe you use the highest-capacity setting because you don't see a visible difference—or maybe you see the difference, but you don't care. I know my car radio/CD player doesn't offer the sound quality of my PC derived-surround-sound speakers or our compact music system—but it sounds great when I'm driving. When I finally decided to try taking music on speaking trips, I picked up an $18 CD player; my wife's comment was, "If you decide you like it, we can get something better." I immediately recognized that the included headphones were atrocious and picked up $10 Sony headphones that sounded a lot better. I did decide I liked it—and I *haven't* gotten something better. For my limited purposes, this $28 combination is good enough.

I don't own an MP3 player, partly because I usually don't listen to music while I'm doing anything else (except driving). I tend to *listen* to music, which interferes with reading, writing, or other high-attention activities. If I did own an MP3 player, I suspect it would be an iPod or competitor and I suspect I'd store music at the highest rate it would accept. I can't imagine watching movies on a 2x3" or 3x4" media-player screen. Even most airplane movie screens strike me as giving up too much of the movie's detail and quality to be worth watching. But those are my sensitivities; they may not be yours.

A Caution in Closing

Time to bring this meandering and possibly irrelevant essay to a close. I would offer one caution—something you might pay attention to if you believe you don't care about some of these differences.

Do you find that you don't want to listen to your MP3 player for very long—that it becomes tiresome? If so, you may be dealing with compression losses and artifacts at a subconscious level. Try listening to the same music on CD or ripped at a much higher data rate; see if you find the music more involving, more satisfying. It may "sound the same," but you may enjoy it more. Or you might not.

The same goes for compressed TV. Digital artifacts aren't always obvious, but they can be tiring. You might enjoy that convenient video more if you record it at a higher data rate. Or you might not.

Library Access to Scholarship

One of the biggest events in the past few months is the publication of *Scientific publications: Free for all?*, the report of the UK's House of Commons Science and Technology Committee (comments on some of this committee's hearings appeared in the June issue). Notes on that report and early reactions appear in the final section of this roundup. Before that, it's worth noting a few of the many developments in other areas of library access to scholarship—and offering a little perspective.

My primary interest in this section is freeing up library funds so academic libraries can maintain humanities subscriptions, buy monographs, other books, and media, provide access to gray literature, maintain technical services and reference librarian-

ship, and in other ways preserve the record of the civilization and maintain themselves as libraries.

OA journals can help—*if* they're represented in library catalogs and *when* they replace overpriced commercial journals or force those journal publishers to reduce prices. As for OA archives, as far as I can tell, these are likely to have either *no* effect on library costs or—when they have an effect—a potentially disruptive effect on scholarly communication.

As long as OA archives represent such a small percentage of the papers in a given subscription journal that libraries must retain their existing subscriptions, then the OA archives don't help the financial problem at all. When a large enough percentage of the papers in a given journal are represented in OA archives, and the OA archives are harvested so that libraries can reasonably expect to find those papers via OpenURL or otherwise, then a growing number of libraries can, will, and *must* cancel their subscriptions to those journals. That has one effect in the short term, another in the slightly longer term. In the short term, profit-oriented publishers will raise prices for remaining subscribers, squeezing the biggest stones for as much blood as possible. In the slightly longer term, the subscription journal will fail—*taking with it the full-text archives and the peer-review mechanisms*. The peer review mechanisms will be replaced, of course, as researchers migrate to OA journals. Full text archives may or may not be so easy to replace, unless LOCKSS and national-library archival agreements take care of the situation. The concept that libraries must *and will* retain expensive subscriptions as long as *any* significant papers are being published in those journals that are not available via other means is ludicrous in a world of limited library resources.

The Big Deal

Harry Kriz (Virginia Tech) forwarded an April 6 report from the Roanoke *Times & World News*, "Tech bargains its way to better journal deal." The story recounts the collective bargaining of seven Virginia public university libraries to negotiate a new five-year contract with Elsevier. The local angle is that Virginia Tech receives some 600 additional scholarly journals. The universities recruited a lawyer to assist in the negotiations. Here's the paragraph that startled me at first reading:

> Under the $27 million contract, each library will have access to more than 1,800 Elsevier journals—roughly a 50 percent increase—for a price jump of just 5 percent. The contract also insulates the colleges from future drastic price increases for journals, some of which cost more than $10,000 a year.

$27 million dollars! That's more than twice what the 10-library University of California system was spending with Elsevier before the Big Deal was renegotiated at UC. A little clarification was in order, which Kriz was able to supply from Paul Metz, Virginia Tech's Director of Collections: $27 million is the *total* five-year price, making it $5.4 million a year. That's still a load of money, and there's the assumption that the 600 additional journals are needed and useful—but the group of libraries is spared annual price increases and negotiations.

Society Publishers

Some professional societies and groups representing those societies had comments after the "Free for all" report. Additionally, the odd multipart conversation about how profits from publishing sustain other society activities—and whether that's a plausible economic scenario for the 21st century—continued.

Jan Velterop & European Geophysical Union

Jan Velterop noted a letter in *The Guardian* about a shift by the European Geophysical Union to publish its journals with full open access. "This is not only to be enthusiastically welcomed by the scientific community, but also a very strong indication that publishing with open access is *not* beyond the possibilities of learned societies with the political will to do so, in spite of frequent assertions to the contrary." He also invited societies nervous about the hassles of going the open access route to contact BioMed Central "with a view to sharing [its] tried and tested tools and platform…without the need for any investment up-front."

Velterop took issue with two statements by William Sturges, the letter's author. Sturges is "not quite correct…in stating that open access is 'driven' by learned bodies 'with for-profit publishers running to catch up.' If only. Neither societies nor commercial publishers are as groups by definition on one side of the divide." He reminds us that open access publishing on a serious scale was first done "by a commercial (i.e. unsubsidised) publisher: BioMed Central. Most for-profit publishers are not exactly 'running to catch up' and most societies are not (yet) driving the process."

You could argue that "unsubsidized" and "commercial" are two very different animals. Certainly under U.S. law, most nonprofit organizations are unsubsidized. And while BioMed Central may be commercial, it's hard to call it a for-profit enterprise until it manages to turn a profit. I have no such nits to pick with Velterop's other disagreement with Sturges' letter:

> Another statement in the letter is unfortunately not quite correct, either, although, again, may be correct for the earth sciences. That is that the EGU journal would be the first "truly" open access one in that the

peer-review process is open and published as well. In the medical sphere, BioMed Central's journals as well as the British Medical Journal (BMJ) have been operating along these lines for years.

Blackwell and ALPSP

According to a Blackwell newsletter for July 2004, Blackwell carried out a survey with ALPSP on what societies do with their publishing surpluses. Sixty-eight societies responded. Most respondents (all but three) subsidize member access to journals—either free or at a reduced price. Surpluses also help cover general expenses, are reinvested in publishing, help reduce conference fees, and cover other activities.

Societies that do their own publishing clear an average 18% surplus from journal subscriptions, covering 32% of *total* society income. Societies using commercial publishers earn surpluses averaging 33% of society income.

Here's the final paragraph of the article:

> In the current debate on Open Access, the importance of journal profits to societies has been mentioned and indeed the right to make such profits has sometimes been questioned. This survey shows the significance of journals in the finances of societies and the benefits to members who use the journals.

David Goodman forwarded that paragraph to the SPARC Open Access Forum, adding this comment: "I would add a final sentence: Whether it is libraries who should fund these activities, is another question." I sent him a note saying I planned to add a longer version of that response, and I've said the same in the past—that is, it is *not* reasonable for libraries to be expected to subsidize the activities of professional societies other than those for libraries—but there's not much more to say. If it's reasonable for universities to subsidize professional societies, that subsidy should be direct—not indirectly through overpriced journal subscriptions.

After writing these comments, I downloaded and read the survey report itself, "What do societies do with their publishing surpluses?" The nine-page report provides more detail on the numbers—for example, one-third of the respondents said they *don't* make a surplus on their publishing—but doesn't require much more commentary. It's a descriptive article, not a combative one. In fairness, and since the newsletter piece did not raise the question Goodman wants to raise, I should quote the last paragraph of the executive summary, following the list of society interests that would not be as well served if there were no surpluses:

> Whether library budgets (both academic—i.e. the taxpayer—and industrial) are the best way of serving these interests is a question which is currently being hotly debated.

American Society for Biochemistry and Molecular Biology (ASBMB)

In another contribution to *Nature*'s ongoing "web focus" on access to the literature, Bettie Sue Masters (University of Texas Health Science Center at San Antonio) and Judith S. Bond (College of Medicine, The Pennsylvania State University) commented on ASBMB's practice.

The flagship journal is the *Journal of Biological Chemistry* (JBC), published since 1905. (The society is a year younger than the journal.) ASBMB now has 12,000 members and also publishes a monthly magazine and three other journals. *JBC* became available electronically in 1995, "the first biomedical journal to be available electronically" in conjunction with HighWire Press. "In 2001, *JBC* introduced Papers In Press (PIPs), which makes manuscripts available online the day they are accepted for publication, and permits free access to *JBC* papers to anyone." Accepted articles typically appear online around 8 weeks before they appear in print—and *stay* available online. *JBC* now provides "free, online, full text searchable access to every published article since its inception in 1905," and the sister journal *Journal of Lipid Research* also provides free online access to every article since its founding in 1959. Masters and Bond say ASBMB spent more than $700,000 to make this happen—made possible by the multiform business model of the journals. "Our expenses are paid by a combination of sources, primarily page charges to authors and subscriptions to individuals and libraries. In a recent survey of over a 1,000 *JBC* authors, over 80% preferred this mode of covering expenses to other models, such as authors or institutions paying all the costs."

Open access works for ASBMB. Submissions continue to increase, as do citations. ASBMB maintains a significant full-time staff to support its publications. The essay includes a seemingly obligatory swipe at pure "author-pays" models, an unfortunate distraction from the rest of the discussion:

> The subject of editorial independence cannot be ignored. Depending upon the business model, unless large submission charges are levied, there may be a tendency to lower the standards of peer review to permit more manuscripts to be published. There is risk, for example, in an author-pays-all-costs publication model that standards could be influenced by the acceptance rate of manuscripts.

That's followed by a statement that a high standard is the creed for society publishers—but that's equally true for any publisher who expects to build a reputation for its journals. I don't disagree with a later statement: "As any believer in the free enterprise system would espouse, it is better to allow and, indeed, to encourage competition among various modes of

publication." But including the standard canard about lowered review standards is an unfortunate and, I would think, unnecessary part of such competition—*particularly* for a society that appears to be providing open access.

Open Access Publishing

BMJ announced 2005 charges—but ones that don't affect OA status. To wit, full text of *non-research* items (Editorials, Reviews, Letters, etc.) will require subscriptions for online access from the 2nd through 52nd week of publication. All content less than a week or more than a year old will continue to be available for free online; original research articles will continue to be free; abstracts and extract views will continue to be free; and more than 100 developing countries will have full free access. This change seems consistent with OA principles: Refereed articles are freely available, with added value for a fee.

Cell Communication and Signaling published a brief article to review the reasons that the journal's publisher is introducing article-processing charges (APC). Since the fee ($525) is levied only for *accepted* articles, this places *CCS* in the same business model as other BioMed Central journals. It's a good brief review of the advantages of Open Access: www.biosignaling.com/content/2/1/7

The *Proceedings of the National Academy of Science* announced the same hybrid option (or Prosser model) some other publishers have chosen, as an experiment through the end of 2005. If PNAS authors pay a $1,000 surcharge, their articles will be available for free via PNAS Online and PubMed Central immediately upon publication. At the end of the experiment, PNAS may continue as a hybrid publication, move toward full Open Access, or retreat from the option. The press release notes that PNAS "operates as a nonprofit, break-even operation." (Authors from institutions with 2005 PNAS institutional site licenses will pay a $750 surcharge.)

Swets Information Services announced in early July that it would add BioMed Central's Open Access journals to SwetsWise Online Content, which effectively makes the already-free full-text articles more conveniently available for some institutions.

Springer has announced its version of the Prosser model, which it calls "Open Choice," for *all* of its journals—but at a price: The author charge is $3,000 (plus possible page charges for print editions), twice that of PLoS and almost six times that of BioMed Central. Springer's CEO, Derk Haank, was formerly CEO of Elsevier. Jan Velterop was quick to note that the Prosser model is *not* Open Access—and that Springer continues to require a range of copyright and license agreements that make article use much more restrictive than in OA journals. An article in *Financial Times* (excerpted by Peter Suber) suggests just how sincere Springer may be. "A vocal minority of libraries and academics are also calling for a revamp of the traditional 'user pays' publishing model, which they claim is too costly for the end user. Instead, some are promoting a so-called open access model in which an author or sponsoring institution pays to have articles published that are then widely disseminated. Mr. Haank says the debate, which has pitted some open-access upstarts against the industry leaders, has taken on an 'unhelpful,' 'almost religious' emotional element… One rival says Springer's plan represents little more than a 'public relations initiative.' It is an accusation Mr. Haank would likely deny, although he does appear to relish the challenge he is presenting to some academics to put their money where their protest are. 'Let's see how serious they really are…we expect that not more than 10 percent will be interested in this option,' he says." As Suber notes, "Haank sounds as if his plan is designed more to generate low uptake, and ground a rebuke to OA advocates, than to test the waters in good faith." Sure sounds that way to this interested observer. "User pays" may be an even more misleading name for the current state of STM journals than "author pays" is for one OA publishing model: It's *not* the users (primarily scientists) who pay, but the libraries.

Oxford University Press introduced a new OA journal, *Evidence-based Complementary and Alternative Medicine*, with an unusual support mechanism: For the first decade, Japan's Ishikawa Natural Medicinal Products Research Center will support the journal (which "will focus on traditional Asian healing systems"). OUP also announced that *Nucleic Acids Research* would move to full OA publishing beginning January 2005, with mandatory publication fees and immediate free access to articles. NAR is a high-impact journal and has been around more than three decades; the contribution fee is $1,500.

The director of the National Institutes of Health suggested at a late July 2004 gathering of journal publishers and editors that eventually, all NIH-financed research must be freely available to the public. Unsurprisingly, reactions included one from an AAP officer, who said that NIH's recommendations could undermine the sustainability of the publishing industry and exert a "chilling effect" on NIH-funded authors. Barbara Meredith, quoted in a July 21 piece at the-scientist.com, said the AAP does not oppose open access, "but it does oppose the government's decision to interfere with the free market by deciding how research should be published." Apparently, the NIH recommendation really isn't for full OA: It allows for a six-month embargo. Peter

Suber had a good comment, noting that NIH is the largest science funder within the U.S. federal government: "The NIH does not work for the publishers. It works for the taxpayers." At least that's the theory. Reactions also included support from the Genetic Alliance, a coalition of organizations advocating for patient families.

SLA issued an odd "statement regarding open access" on June 5, noting that "SLA has traditionally supported accessible information, but not necessarily free." The statement "encourages ongoing exploration of viable means to expand the availability of scientific and scholarly research," but falls far short of endorsing OA. The British Columbia Library Association *did* formally endorse OA in a June 19 resolution—including a statement encouraging libraries to support OA by cataloging and providing access to OA journals, a necessary additional step towards making OA effective.

Brief Commentaries

Hodgkin, Adam, "After the tipping point—what next?" May 14, 2004.

Hodgkin is president and co-founder of xrefer, a "leading reference aggregator." The "tipping point" he's discussing is the next one after the one that starts to bring down the big deal. He quotes John Cox from the *Charleston Adviser:*

> One thing I do know is that a number of major commercial publishers are involved in contingency planning if Open Access reaches the 'tipping point' at which the whole industry switches business models. Open Access will not lead to the demise of the large commercial publishers…

Hodgkin agrees with Cox's implied prediction, that is, when this tipping point is reached there will be a *decisive* industry-wide switch to an OA model. I'm not nearly as convinced, but I'm not part of the industry. Hodgkin thinks the decisive moment will be "marked and accelerated" when one or two big publishers endorse and adopt the OA business model. "Scientific publishers respond to the needs of science and at some point the publishers, not the contributors/authors, will do the decisive tipping." (I'd guess the biggest publishers respond primarily to the needs of stockholders, but never mind…)

He notes the advantage of being an early mover in OA if it looks likely to become a favored model; that there "will be considerable economies of scale" for players with the right infrastructure (although it's unclear that an effective OA platform *requires* huge scale to be economical); that the costs of processing papers could be pushed below $100 per paper; and that other parts of scientific publishing *could* suffer if OA takes over—but not necessarily. He goes on to posit that efficient OA publishing could "encourage the publication of yet more papers" and thus increase the overload of primary research publishing. This might open up new opportunities for secondary and tertiary publishing—review publications, survey and background periodicals, major reference works with synthetic coverage, specialist databases and the like. His conclusion, which strikes me as perfectly reasonable:

"Publishers will not enjoy or acquiesce in losing their most profitable cash cow journals; but there are reasons for thinking that an open access world for primary research will still leave plenty of scope for profit-seeking and innovative scientific publishing."

George Porter: Two Commentaries at STLQ

Porter offered two brief, cogent commentaries on two different aspects of access to scholarly communications on May 14 and May 26, 2004; you can find them at stlq.info/archives. The first, "The crisis in scholarly communication," notes that the crisis is into its third decade—but recent years show a change in awareness and concern. Porter's been "tracking the ripples" caused by the resignation of Donald Knuth and the rest of the editors of *Journal of Algorithms* (an Elsevier publication) and the ensuing launch of ACM *Transactions on Algorithms* (discussed here in the March 2004 issue). He sees one ripple in a *Stanford Magazine* article on the skyrocketing costs of scholarly journals and another in the *Report of the Seventh meeting* of the Committee on Electronic Information and Communication (CEIC) of the International Mathematical Union.

The May 26 commentary deals with a different set of issues: "When a journal ceases publication." Porter asks the question, "What happens when a journal ceases to produce new issues?" For print, the received collection stays in place unless librarians decide to get rid of it—control is entirely local.

For ejournals and those only held as etext, however, the choices aren't as clear. (Porter notes that some STM publishers play fast and loose with ISSN standards, making it harder to track journals.) When a print title changes publishers, libraries barely notice (and don't usually update cataloging records)—but if an ejournal changes publishers, "volumes may disappear, be transferred, have redirects, or a number of other variations."

Porter offers distressing examples. Springer Verlag published *Nonlinear Science Today* beginning in 1992, originally in print only. After becoming a pioneer in e-publishing, the journal dropped its print version in 1994—and ran into trouble, with four issues stretching over 1994 and 1995 and a single-issue "volume 6" in 1996. The articles continued to be available online—until Springer migrated to the

SpringerLink MetaPress platform in July 2003. At that point, the website vanished; articles have apparently disappeared entirely. "Springer's journal title list no longer acknowledges the former existence of the title. It's simply gone, without explanation, without a trace." Several other cases appear to be happier, at least so far.

> Authors, publishers, and librarians are philosophically united, I trust, on the value of the intellectual record. This is definitely broader than simply ceased journals, but the solutions, which address the broader issue, may have a dramatic impact on the narrower.

Porter notes LOCKSS (see January, July, and August 2003 issues) as one partial solution. It has entered production as of April. A number of important publishers are participating, including Oxford University Press, Kluwer, Blackwell, Nature, and the ones you'd expect—but not Wiley, Elsevier, Taylor & Francis, IEEE, ACM, and a number of other important professional societies.

> Time will tell if/when more publishers will discover enlightened self-interest. The prospects for growth in library participation are significant. I am less sanguine for the prospects of growth of publisher participation.

Filman, Robert E., "Not free, but relatively inexpensive," *IEEE Internet Computing* July/August 2004.

This editorial covers a lot of ground—belittling most web content, noting that most labor involved in academic publishing is unpaid, and noting that *printed* journals have real costs. Then Filman asks, "What keeps academics from just publishing on the Internet? Very little. Purely Internet journals are springing up." Publishers recognize the need to make journals available on the web—but that's constrained such factors as "the actual editing services offered by revenue-generating publications." He seems to be posing a dichotomy: *either* traditional journals *or* internet publications with no revenue stream and, presumably, no editorial work.

"Academic publishers do add value." Who doubts that? Filman contrasts this with "a world rushing toward the most economical way of doing things" and a population that "has come to believe that…information on the Web is naturally free." Later: "The continued existence of the formal, paper-based, peer-reviewed publication depends on the (primarily University-based) research community's resisting the powerful economic forces. Unfortunately, I bet society will take the less-expensive road in the long term."

Maybe I'm reading too much into this editorial. Maybe Filman isn't indirectly attacking ejournals and the OA model by directly coupling peer review with traditional publishing and failing to note alternative financial models. Maybe he isn't implying that, if libraries stop paying whatever traditional academic publishers want, the result will be a loss of editing and quality. But that sure is the way this piece comes across.

Cockerill, Matthew J., "Delayed impact: ISI's citation tracking choices are keeping scientists in the dark," *BMC Bioinformatics* 5:93 (July 12, 2004).

This editorial expands on the statement in the title, with examples from BioMed Central, one of the premier early OA publishers. To wit, ISI's "impact factor" is heavily used as a measure of STM journal quality—but no impact factor is available until the third year after ISI starts tracking a journal. (The impact factor is "calculated by dividing the number of current citations to articles published in the two previous years by the total number of articles published in the two previous years.") ISI doesn't track everything, which makes things worse.

This journal is used as an example. It started publishing in 2000 but ISI didn't start tracking until 2002, so the first impact factor (for 2004) won't appear until June 2005. But it's possible to prepare an "unofficial impact factor" by using ISI's cited reference database, which includes all references in tracked journal articles. Using that methodology, the 2003 impact factor for *BMC Bioinformatics* would be roughly 4.9 (235 citations for 48 articles), which would place it in the top 5% of journals covered by ISI. A new author wouldn't know that: The journal simply doesn't appear in the 2003 *Journal Citation Report*. This situation with many newer journals may dissuade authors from publishing in those journals.

Cockerill expresses the hope that competition in citation analysis may encourage ISI to "reconsider its policy on citation tracking," immediately tracking any peer-reviewed journal that meets basic quality standards and can provide reference list data in an appropriate form for automated analysis. "By doing this, ISI would provide a valuable impartial service to the scientific community."

"Open Access Journals: revenue beyond author charges," Yale University Science Libraries, August 9, 2004. www.library.yale.edu/science/oa.html

This is a new web page that attempts "to outline the current pricing models that are being tested for supporting Open Access to electronic journals." The page links to a number of related pages. The author—presumably David E. Stern (maintainer of the

page), although it's not signed as such—says, "The major concern of the community should be maintaining a revenue stream to support the peer review process." He goes on to discuss varieties of OA (and near-OA) in existence today, the nature and problems with author charges, and some alternatives.

If you're thinking about revenue issues and the possibly disruptive effects of OAI and OA journals (see my editorializing at the top), I **recommend** reading this page—it's just over three pages—and thinking about it. I'm not saying I agree with the summary; I'm saying this is a thoughtful discussion that leaves me feeling that I don't know enough to have a useful opinion.

Longer Articles and Commentaries

Suber, Peter, "Open access overview," June 21, 2004. www.earlham.edu/~peters/fos/overview.htm

If you want a clear, concise understanding of what open access is all about, **read this introduction**. Better yet, offer it to others who either don't understand or misunderstand open access. It's clear, brief (seven pages of bullet points), and seems to cover the territory. A few quotes, sometimes slightly out of context (occasionally combining portions of bullet points):

> The legal basis of OA is either the consent of the copyright-holder or the public domain, usually the former. Because OA uses copyright-holder consent, or the expiration of copyright, it does not require the abolition, reform, or infringement of copyright.
>
> OA literature is not free to produce or publish. No serious OA advocate has ever said that OA literature is costless to produce, although many argue that it is much less expensive to produce than conventionally published literature, even than online-only toll-access literature.
>
> OA is compatible with priced add-ons.
>
> OA is compatible with peer review… Removing access barriers and reforming peer review are independent projects. OA doesn't presuppose any particular model of peer review…
>
> The chief difference between [OA journals and OA archives or repositories] is that OA journals conduct peer review and OA archives do not.
>
> A common misunderstanding is that OA journals all use the "author pays" business model.
>
> We can be confident that OA journals are economically sustainable because the true costs of peer review, manuscript preparation, and OA dissemination are considerably less than the prices we currently pay for subscription-based journals. There's more than enough money already committed to the journal-support system. Moreover, as OA spreads libraries will realize large savings from the conversion, cancellation, or demise of subscription-based journals.#
>
> OA archives are economically sustainable because they are inexpensive.#
>
> OA is a kind of access, not a business model.
>
> Open access is not synonymous with universal access.
>
> *Libraries.* OA solves the pricing crisis for scholarly journals…#

I could raise mild questions about the seventh, eighth, and last of those excerpts (each marked with a # at the end of the quoted section):

- While I'm as confident as Peter Suber that OA journals are economically sustainable (if only because some of them have been around for quite a long time), the argument here supposes that money spent on subscriptions will become available for OA funding mechanisms. That begs the "freeloader" question—those private enterprises that never publish research in scholarly journals but make heavy use of such journals, and schools that have few publishing scholars but subscribe to many journals. The final sentence is what I hope for—but it's *only* directly related to OA journals, not the full scope of OA, and there may be a substantial problem getting from here to there.
- Yes, OA archives are economically sustainable—but are they *politically* sustainable? That depends on making them integral parts of the ongoing academic missions of the institutions. If an institution closes down a department (which does happen), what motivates it to keep that department's OA archive active? (This argues for making libraries the political center of institutional archives: No academic institution worthy of the name will close down its library, although many have been systematically underfunding their libraries.)
- OA in and of itself does *not* solve the pricing crisis for scholarly journals. OA archiving does nothing at all to alleviate the crisis. OA journals may help, *to the extent that they replace or force reductions in the price of commercial journals*. Otherwise, OA journals *increase* library costs (albeit slightly): a journal that isn't cataloged and represented in the library's full-text pointers is not "there," no matter how free it may be—and these technical services steps, called "overhead" by some, are not free.

Those nits picked, this is a solid introduction and, I think, about as short as a real introduction to OA

can be. Suber plans to revise it; with luck, you'll see an even better version than the one I downloaded.

King, Donald W., "Should commercial publishers be included in the model for open access through author payment?" *D-Lib Magazine* 10:6 (June 2004).

This is an opinion piece, not a refereed article, a distinction *D-Lib* is careful to make. While King makes the unfortunate simplification of "author payment" for publication-fee schemes, he does say that OA publishing "appears to have real merit and warrants careful examination and testing—a view I have not always held. However, I also believe it is counter-productive for author payment advocates to denigrate commercial publishers (and profit) and exclude them from the open access model. To do so diminishes the chance of success for the model in the long run."

Obviously I've been missing something in the OA discussions. Peter Suber certainly doesn't argue that for-profit publishers should be excluded. Jan Velterop considers BioMed Central to *be* a for-profit publisher. I don't remember seeing this argument from PLoS (although I haven't read all of that group's position papers). Loath as I am to *ever* raise the issue of straw men, I wonder just which advocates King is talking about?

"The commercial sector has made too many contributions to science to dismiss the sector as being irrelevant and basically self-serving." I've certainly never seen a claim that commercial publishers are "irrelevant," although "self-serving" strikes me as a fairly accurate description of most of them. There's nothing wrong with being self-serving as long as you serve others as well.

Once we get through some historical comments, we get to a discussion of actual article processing costs that I find questionable. King asserts that an article processing cost of $3,000 per article is "not atypical for traditional science publishers," later says that traditional publishers indicate a cost of "from $2,000 to $4,000 with all costs included," and offers this expert opinion: "I believe that it may typically be in the $3,000 to $4,000 range…"

Based on what little evidence I've seen (some of it mentioned in previous issues), "$3,000 to $4,000" is only plausible as a "cost" range if you include as part of "article processing" all corporate overhead, all sales offices, all *current* profits, all costs of maintaining subscriptions, and all current corporate salaries: In other words, everything except printing and mailing. In other words, $3,000 to $4,000 represents the high end of per-article *revenue* (except for megajournals like *Science)*—and even there, Blackwell (for example) claims a much lower figure. PLoS's tables suggest that $1,500 is on the high side for article processing with a 90% rejection rate; BioMed Central expects to become profitable at $525 per article.

King also says, "The investment necessary to replace an existing commercial journal tends to be about $100,000 for start-up, capital requirements, future research and development, and operations. Thus total investment to replace all commercial journals would be on the order of hundreds of millions of dollars…" Here again, I think you need more evidence that an e-journal with no subscription handling, sales force, or other overhead needs such a large startup fund, particularly given the growing numbers of inexpensive shared platforms for manuscript control and publishing workflow. Some established players are only too happy to work with new OA journal publishers to minimize startup costs.

I don't know whether to recommend this or not. I agree that commercial publishers should be involved in OA experiments and conversions (and so do most OA advocates that I know of). I don't believe King's asserted *cost* for article processing has been or, in fact, can be demonstrated: It's implausibly high and assumes that today's high profit margins, huge corporate overhead, massive subsidies for non-publishing activities of society publishers, and costs of maintaining subscriptions and restricted electronic access should all be maintained in an OA world. Maybe I'm reading it wrong.

Johnson, Richard K., "Open access: Unlocking the value of scientific research," presented at *The new challenge for research libraries: Collection management and strategic access to digital resources*, a conference sponsored by the University of Oklahoma, March 4-5, 2004.

While this presentation carries a somewhat-unnecessary copyright notice, it also carries a Creative Commons Attribution License note, allowing unrestricted use with proper citation: You could use this in a for-profit book without asking Johnson.

It's a good piece from a SPARC official, **well worth reading** for a library perspective on OA. Johnson points out that the STM pricing crisis has been going on "for as long as most of us can remember" and that, in the print world, "this was seen mainly as a library issue," largely invisible to faculty. It's becoming a broader issue. Quoting Ross Atkinson: "This is not a serials crisis, but rather a broader crisis in scholarly communications."

Johnson notes Big Deals and increasing library resistance to such lock-in forces, including a Goldman Sachs survey that nearly a quarter of librarians

planned to cancel or reduce Elsevier ScienceDirect subscriptions, with another third demanding price cuts. He notes the extent to which, paradoxically, the largest publishers can actually *gain* market share by increasing prices, as libraries are forced to cancel other subscriptions to retain the top journals from the biggest players.

Johnson calls OA "a scalable solution that addresses the economic dilemma of libraries at the same time as it exploits the potential of the networked environment." He refers to OA as "an outcome that may be supported in a range of ways with an infinite variety of business models." A table from J. Willinsky shows not two but *nine* types of OA, although not all of them represent true open access. He quotes Alastair Dryburgh on the likelihood that the high-profit commercial publishers will be early adopters of OA journal publishing: "They may in fact need to be dragged kicking and screaming into the new world." Dryburgh thinks this could happen either because funding bodies *insist* on OA or because effective harvesting of OA archives for a significant proportion of the literature leads to the kind of disruptive changes suggested in my editorial comments ("subscription attrition will turn into a rout and open access will become the only viable model for the publication of primary research").

Johnson says—and Peter Suber also notes, below—that the Sabo bill never progressed, "but it sent shock waves through the scientific publishing industry." That's probably good on both counts. Discussing scholarly societies and OA, Johnson notes, "many societies…are profoundly skeptical of open access… Typically they fear the disappearance of surpluses from institutional subscriptions that support other activities of the society." He goes on to discuss the American Society for Cell Biology, which runs a profitable annual conference and is moving toward open access. How fast will societies move? One major society has a white paper entitled "Open Access to [society's name] Publications by 2020?" That's all deliberate speed with a vengeance!

> Perhaps the greatest obstacles to open access today are: the risk that journal publishers will not recover sufficient revenue to cover their publishing costs or generate a sufficient surplus; and lack of author awareness of the benefits of depositing their work in open access repositories.
>
> Libraries and their institutions are in a position to do something about these obstacles. For example, they can: Establish institutional repositories; Help faculty archive their research papers…; Help open access journals…become known…; Insure that scholars…know how to find open access journals and archives…; As [OA journals] proliferate…libraries can cancel over-priced journals…; Engage funding bodies in a discussion of [OA]; Familiarize faculty, staff, and administrators with the issues. *[Note: These clauses are shorter versions of bullet points in the original.]*

Johnson concludes, in part: "The essence of the case for open access is the notion that the public good—the societal benefits derived of our research investment—is better served when barriers to sharing of research have been removed. That belief aligns well with library values." Indeed.

SPARC Open Access Newsletter 74, 75, 76 (June 2, July 2, August 2, 2004).

As always, if you're interested in OA, you should be getting *SOAN*, so I'll point out just a few highlights from these three issues.

June's issue focuses on Elsevier's postprint archiving permissions (which are tricky but may still be a "breakthrough"), asserts that authors are the key players in making OA happen and suggests "author-centric strategies for achieving OA," and proposes "unbinding projects" to provide retrospective open access to the key research articles on various topics, based on authoritative bibliographies.

In July, Suber continues to discuss Elsevier's new policy and some of its implications—and here notes that "OA to the literature" in the form of archiving isn't inherently "the kind of OA that helps libraries." (Am I getting through, or is Suber hearing this grump from others as well?) That lengthy discussion is the only long essay in the issue, which as usual includes summaries of major OA developments and links to relevant articles.

August includes lead stories on "two of the most significant open-access developments in our history. It's uncanny how similar they are and how, without planning, they were announced in the same week, reinforcing each other's message and momentum."

The second development is the UK House of Commons report discussed below. The first, which had slipped entirely under my radar ("NIH" doesn't even appear in my running index for this volume), puts some teeth behind the NIH comments noted earlier. To wit, the U.S. House Appropriations Committee has adopted a recommendation for next year's federal budget that would have NIH put a condition on its research grants: Articles based on NIH-funded research would be deposited in the freely-available PubMed Central six months after publication—and if NIH paid *any part of* the publication costs, the article would be deposited (and freely available) *immediately* upon publication. The first part is a baby step toward OA, establishing a limit to publishers' embargo periods; the second is a huge step forward if it survives the legislative process.

Suber offers "ten annotations to help understand the proposal." The annotations are **well**

worth reading (a constant reminder for *SOAN*, but worth repeating in this instance). The ninth and tenth deal with some early and seemingly inevitable attacks on the proposal—first, that it diverts funds from NIH's research mission; second, that we should "let the market work." The diversion of funds argument—from the executive director of the American Physiological Society—includes an estimate that the new scheme would cost NIH $75 to $100 million (per year?); an estimate labeled "too high by at least an order of magnitude" by the director of PubMed Central. Suber's response to the "let the market work" objection is one I consider nearly unassailable: "Almost every observer not paid by publishers believes that the journal publishing system is dysfunctional and unsustainable… Insofar as it's a market, it has failed." And, to be sure, this calls for mandated OA archiving, *not* OA publishing; by itself, it does little to undermine overpriced journals.

Suber goes on to contrast this plan with the unfortunate Sabo bill: "Procedurally it's still alive but politically it's been dead for some time… It died because it was written so that even OA proponents could not line up behind it." The NIH proposal does not affect copyright law, doesn't accidentally include materials that need to be excluded, and appears to be a nonpartisan or bipartisan effort. Suber calls it "an extraordinary step forward."

In summarizing the UK report, Suber compares the U.S. and UK developments; it's a good comparison that I won't bother to repeat.

Esposito, Joseph J., "The devil you don't know: The unexpected future of Open Access publishing," *First Monday* 9:8 (August 2004).

You may have seen references to this article in various weblogs. I believe a thoughtful, careful, eloquent, dispassionate evaluation of Esposito's arguments and assertions is warranted. I'm afraid I can't provide it. The forest of red on these 16 pages after I made my preliminary pass seems to rule out dispassionate comment.

A first-pass commentary ran to more than 2,000 words. It was more of a rant than a commentary. I'm not sure Esposito deserves that much space. Doing a little checking, I realized that he also wrote an article about a year ago in *First Monday*—one that, after struggling with a CHEAP SHOT commentary for a while, I finally chose to ignore entirely.

I am *not*, as you might already guess, recommending that you read this as a serious treatment of the likely future of OA publishing or publishing in general. After reading it the first time (and washing my hands afterwards), I set it aside in the hope that someone else would provide the deconstruction this construct so richly deserves. I still hope that.

I find Esposito's treatment sneering; insulting to libraries, researchers, and the public; and singularly devoid of fact or logical argument. He's one of those who appears to see libraries as nothing more than article-pushers; he seems to think that libraries only *license* publications (which will come as a shock to acquisitions librarians); he implies that OA is coupled to loss of copyright; and he dismisses peer review as an artifact of the Gutenberg era. He claims OA advocates assert that it will "reduce costs to zero or some modest mark-up over zero," a claim I've certainly never seen for OA *publishing* advocates. He accuses OA advocates of having the "Change One Thing worldview," an unjustified insult to Peter Suber, George Potter, Jan Velterop, the founders of PLoS and many other thoughtful advocates. He seems to equate OA publishing with weblogs and vanity publishing. He dismisses the usefulness of research papers outside the closed circle with a sentence that deserves direct quotation: "By definition, if someone without sophisticated training (that is, our Man in the Street) could even understand a research paper, then it can't be a research paper."

Esposito's thesis (a dignified term given the sneering tone of much of this paper) is that OA will vastly *increase* the cost of scholarly publishing because publishing's really all about marketing, and once researchers are paying for it themselves, they'll pay oodles more to get more visibility and lots of added services. Or something like that. I'm so angry by the time I get to that part of the paper that I may not be following it properly. **Not recommended**.

Scientific Publications: Free for all?

Volume 1 of this report is impressive—and *long:* 107 pages of small type, the equivalent of a fair-sized book. The conclusions and recommendations come to 82 numbered paragraphs filling 10 pages. Any serious commentary on the publication would take at least ten *Cites & Insights* pages, even assuming I was qualified to do the commentary. Instead, here are just a few interesting points raised within the document, *not* including the conclusions and recommendations (which appear in boldface within the report proper)—not that these are all new points, but they show a little of the considerable depth of the UK committee's understanding of the situation. There are also some warnings scattered throughout—including an explicit note that "author-pays" funding, in the UK at least, is likely to come straight out of library budgets. You might refer back to the June commentaries and note the extent to which the committee took commentaries seriously—or not:

- Peculiarities of the [STM article] market: "The point of purchase is not always the same as the point of use. Libraries purchase journals *on behalf* of their community of users. This characteristic of the market has the effect of insulating readers from the consequences of fluctuations in journal prices… There is a lack of substitutability in the market…"
- "We understand that many journal articles are esoteric… Nonetheless, we cannot see what damage could be done by allowing the public to examine the articles for themselves. Unlike Dr Jarvis, the possibility of better-informed patients 'marching into surgeries and asking things' does not fill us with horror."
- "Rising STM journal prices also have an impact on the library's provision of other information."
- "There is widespread discontent amongst libraries with bundling."
- "Usage does not equate to usefulness. Niche journals publish research of minority interest that is nonetheless of great importance to those who work in the field."
- "It has been argued that public money is used at three stages in the publishing process: to fund the research project; to pay the salaries of academics who carry out peer review for no extra payment; and to fund libraries to purchase scientific publications."
- "The rate of £11 ($20) per article reviewed strikes us as an acceptable basis for an analysis of costs."
- [Credit Suisse First Boston]: "We estimate that printing and distribution costs are 15% of total costs."
- In the UK, printed publications are exempt from 17.5% Value Added Tax. Digital publications are not.
- "We found it worrying that academics did not take an interest in what happens to their research after it has been published…. Academics have no financial incentives to self-archive."
- "We suspect that the costs per article of author-pays publishing supplied to us by commercial publishers are exaggerated."
- "Any transition to an author-pays model would entail the transfer of some of the library's funds to the research funders to enable them to meet publication costs."
- "If learned societies are valued by their communities, which we believe to be the case, members are likely to remain loyal irrespective of the publishing model employed by their society."
- "Provision of STM journals in the UK is unsatisfactory."

The Recommendations

Peter Suber's summary of the 82 recommendations runs to 12 pages as printed out from his SOAF posting. What follows is further excerpted from Suber's July 19, 2004 posting (with the first part modified in the August 2, 2004 *SPARC Open Access Newsletter)*. I'm leaving out *most* of the recommendations (and particularly most UK-specific recommendations) in the interests of space.

> Here's my summary of the major recommendations:
>
> 1. The government should provide funds for all UK universities to launch open-access institutional repositories.
>
> 2. Government funding agencies should require faculty receiving research grants to deposit copies of their articles in their institutional repositories.
>
> 3. The government should create a fund to help authors pay the processing fees charged by open-access journals. The committee is not yet ready to endorse the upfront funding model for OA journals (which it unfortunately calls the "author-pays" model), but wants to create such a fund in order to promote further experimentation with the model.
>
> 4. The government should develop a wider, long-term strategy that includes open-access journals "as a matter of urgency."
>
> 5. Journal prices are unacceptably high and publisher justifications for them are not credible. The Office of Fair Trading (the UK office investigating monopolistic business practices) should monitor the journal publishing industry and issue biennial public reports on the "state of the market."
>
> 6. The government should investigate whether leaving copyright in the hands of authors would have a "disproportionately negative impact" on authors or research. If it would not, then government funding agencies should require their grantees to retain copyright in articles based on funded research.
>
> 7. All these steps can and should be undertaken without jeopardizing "rigorous and independent peer review."
>
> 8. The government should fund the British Library to take on the long-term preservation of digital scholarship.

A few excerpts from the conclusions and recommendations:

> 8. All researchers, regardless of the nature of their institution, should be granted access to the scientific journals they need to carry out their work effectively.
>
> 10. Teaching is a crucial university function. Universities should be permitted, within reason, to derive maximum value from the digital journals to which

they subscribe by using them for legitimate teaching purposes. We recommend that future licensing deals negotiated by theJoint Information Systems Committee explicitly include provisions to enable journal articles, whether print or digital, to be used for teaching purposes.

11. It is not for either publishers or academics to decide who should, and who should not, be allowed to read scientific journal articles. We are encouraged by the growing interest in research findings shown by the public. It is in society's interest that public understanding of science should increase. Increased public access to research findings should beencouraged by publishers, academics and Government alike.

12. We are not convinced that journal articles are consistently available to members of the public through public libraries.

13. Digitisation should facilitate, not restrict access.

16. We recommend that the Joint Information Systems Committee develop an independent set of measures, agreed by subscribers and publishers alike, to monitor trends in journal pricing. This will help exert pressure on the publishing industry to self–regulate more effectively and will give libraries and other users greater knowledge when they are deciding which subscriptions to take.

17. It is not for us to pronounce on the acceptability of the profit margins secured by private sector companies. Nonetheless, high publisher profit margins need to be set against the context of faltering library budgets and an impending crisis in STM journals provision…

20. Increasing usage rates do not equate to an increased ability for libraries to pay for journal bundles. The recent availability of usage statistics should not be used as a justification for publishers to raise their prices.

21. Although libraries may aspire to provide access to every scientific journal, they cannot afford to do this. It is inevitable that difficult choices between a number of journals with lower usage rates and impact factors will have to be made. Nonetheless, these decisions should be made in response to local user needs rather than as a side effect of bundling.

22. Current levels of flexibility within the journal bundle do not present libraries with value for money…

24. We do not doubt the central importance of peer review to the STM publishing process. Nonetheless, we note a tendency for publishers to inflate the cost to them of peer review in order to justify charging high prices…

26. We are persuaded that the costs to publishers associated with digitisation will reduce over time. Consequently, we would no longer expect these costs to be used as a justification for steep increases in prices…

37. Pressure on library journal acquisitions budgets has resulted in cancelled subscriptions and has contributed to a decline in book purchasing. This compromises the library's ability to provide the full range of services required by its user community.

38. There is undoubtedly some scope for libraries to make efficiency savings, as there is for most organisations. Nonetheless, the valuable services provided by the library are expensive and staff–intensive. It is unlikely that libraries will have more to spend on acquisitions until they see an increase in budgets.

41. It is disappointing that many academics are content to ignore the significant difficulties faced by libraries. Until they start to see the provision of journals as, in part, their problem, the situation will not improve.

42. Elsevier is no sudden convert to Open Access. The company has seen the direction of trends in publishing and has acted accordingly to minimise criticism of its current policies. We are in little doubt that Elsevier timed the announcement of its new policy on self–archiving to pre–empt the publication of this Report...

43. Institutions need an incentive to set up repositories…

47. Institutional repositories should accept for archiving articles based on negative results, even when publication of the article in a journal is unlikely. This accumulated body of material would be a useful resource for the scientific community…

54. Peer review is a key element in the publishing process and should be a pillar of institutional repositories...

58. We see institutional repositories as operating alongside the publishing industry. In the immediate term they will enable readers to gain free access to journal articles whilst the publishing industry experiments with new publishing models, such as the author–pays model.

60. The evidence produced so far suggests that the author–pays model could be viable. We recommend that Government mobilise the different interest groups to support a comprehensive independent study into the costs associated with author–pays publishing…

62. Although early indications are positive, it is too early to assess the impact that author–pays publishing has had on access to scientific publications.

66. In order to succeed, most author–pays publishers, like everyone else, will have to publish articles of a high quality. It is not, therefore, within the interest of journals at the higher end of the market to lessen the rigour of peer review. Nonetheless, there is a risk that lower quality journals might seek to reduce their quality threshold in order to generate profit. Were the author–pays publishing model to prevail it would be vital to ensure that peer review was not compromised in order to retain confidence in the integrity of the publishing process.

67. The introduction of a submission fee would be an important step towards ensuring the quality of

scientific publications and we strongly recommend that author–pays publishers introduce this system.

73. We are satisfied that, by scaling publication with research costs, the author–pays publishing model would ensure a fairer global distribution of the costs of publishing research findings.

75. Institutional repositories should be a key component of any long–term strategy to ensure the preservation of digital publications.

76. The British Library has a crucial role to play in the preservation of digital publications, both strategically and practically. This is an expensive process…

82. As is the case with any process, peer review is not an infallible system and to a large extent depends on the integrity and competence of the people involved and the degree of editorial oversight and quality assurance of the peer review process itself. Nonetheless we are satisfied that publishers are taking reasonable measures to main high standards of peer review.

Early Comments and Reactions

A flurry of press releases and articles followed the report. BioMed Central, SPARC and PLoS were pleased with the report; some commercial publishers were not. Sir Crispin Davis of Reed Elsevier called it "daft" and asserted that universities did not have the time or inclination to create their own archives, according to *The Guardian*. Davis claimed the academic world wasn't responding to OA—"it's been around five years and its market share is still around one percent." Jan Velterop asserted that OA publishing will be "the most predominant model for scientific research within the next five years," according to an August 11, 2004 Jemima Kiss item at dotJournalism.

The IEE "expressed grave reservations regarding the Science and Technology Committee's call for scientific publishers to move to an open-access publishing model…" IEE's July 21 release cited three "fundamental problems":

> First, such a model inevitably increases the pressure to publish, as the more that is published the greater the revenue. There is therefore a risk of undermining the peer review process leading to a reduction in quality.
>
> Secondly, the charge to authors will be substantial, making it less likely that outstanding work from poorer countries will be published.
>
> Thirdly, large commercial organizations with substantial research activities publish relatively little…but are major purchasers of published material. Under the present model, they pay a fair price for this. However, under the 'author pays' model they will have free access. This will remove income from the system and inevitably increase the charge to authors.

Will there *ever* be a press release questioning OA publishing that doesn't use the first "fundamental problem," one of the myths that will not die—and one that's nicely addressed in the report? Similarly, there seems to be a need to ignore the fact that virtually all OA publishers automatically waive fees for authors from poorer countries. The third issue was specifically addressed in the report—which, notably, does *not* call for scientific publishers to move to OA publishing! My father's an engineer; I've always assumed that engineers cared deeply about facts. I find this press release terribly sad.

The Library Stuff

Bell, Steven J., "End PowerPoint dependency now!" *American Libraries* 35:6 (June/July 2004): 56-9.

I'll admit my first response to this article was "Why didn't *I* write that?" But that's the wrong response. Bell does use PowerPoint (at times), or at least has done so. I almost never do: my OpenURL presentation at OSU (which absolutely required PowerPoint) was the first time I've used it outside RLG in five years! So Bell's in a better position to suggest alternatives to PowerPoint and offer ways to make it less of a crutch when you do use it.

The article begins with a striking contrast:

> I'm convinced that our profession's love affair with PowerPoint is stronger than ever. At the last three library conferences I attended, virtually every presentation by a librarian involved PowerPoint slides.
>
> On the other hand, nearly every keynote presenter or invited speaker (almost always non librarians) made little or no use of PowerPoint. Granted, keynotes differ considerably from research-based presentations, but these speakers connected with their audiences effortlessly.

After noting problems with PowerPoint-based presentations and offering alternatives, Bell offers four excellent suggestions: Keep the number of slides to a minimum (10 per hour!); Avoid over-familiar PowerPoint templates; Unless it's absolutely necessary, spare the audience details about your library; and Resist the urge to supply everyone with a printout of all your slides at the start of the program.

He discusses each of those suggestions. The last is one I find particularly interesting, as I had the experience years ago of speaking to a group that had a full-text handout *before* I gave the speech. There was synchronized flipping of pages and—although much of the speech was impromptu—survey forms indicated that people were unhappy that I was reading the speech. I did the same speech a couple of days later, but this time insisted that handouts not be provided until *after* the speech. This time, there

were no such complaints on the survey forms, I had the audience's attention, and the speech was a lot more fun. If you must use PowerPoint and feel the need to have handouts, distribute them afterwards.

The "10 per hour" guideline is also interesting; I'd love to see speakers limit themselves along those lines. When I was sitting through a recent series of otherwise excellent presentations, I counted PowerPoint slides at a rate from 35 per hour to 60 per hour. I've seen speeches where the slides seemed to have *more* text than the speech itself, although that may have been reader fatigue.

Bell offers good advice. **Read it**; think about it.

Berry, John N., III, "Gale/Library Journal library of the year 2004: San Jose Public Library and San Jose State University Library," *Library Journal* June 15, 2004.

The December 2003 "Crawford Files" was based on the Dr. Martin Luther King, Jr. Library—not as the wave of the future, but as a likely success in a special circumstance. At the time, I hadn't been to the joint-use library, but anticipated a followup column in a year or so after visiting it and talking to the two heads, SJPL director Jane Light and SJSU dean Patricia Breivik. That followup won't happen, at least not as a "Crawford Files," and John N. Berry III has done a first-rate writeup that probably does a better job than I would have. Is the joint-use library a success? That "library of the year" award—the second Jane Light's been involved with—certainly argues that it is.

Berry notes that the award probably should be "Libraries of the Year" this time because the King library is such an interesting (and probably unique) "marriage" rather than merger. There are two library directors; there's a "public collection" and an "academic collection" (both available to all users); and most library departments have two heads. They did settle on a single online system and a single website; there's a single library card and combined reference services. Neither library was very well funded (SJPL's funding isn't bad at roughly $35 per capita, but it's lower on a per-cap basis than many surrounding public libraries), both needed new facilities, and the combined project had critical support from San Jose's mayor and the university president.

Results? Much longer opening hours than either had previously; some 12,000 users a day, reaching a million visitors three or four months earlier than expected; substantially higher citizen satisfaction with the library; rising circulation both in the King library and in the growing set of branch libraries. Here's an odd one: Where master's theses collection wasn't circulated previously, the collection is now shelved in a public space and nearly 1,800 theses have circulated since the building opened.

It wasn't a slam-dunk from the beginning, as locals know. Many SJSU faculty members hated the idea; some doubtless still do. Some San Jose residents believed better branches were more needed than a new Main—but Light, the kind of savvy politician libraries need more of, convinced the various Friends groups that by using redevelopment money for the new Main they could get a bond measure passed to improve branches. It worked: a $211 million bond measure, the first in San Jose in 30 years, passed—enough to build six branches and renovate 14 more.

A good article about an unusual library. **Worth reading**.

Block, Marylaine, "On analyzing web sites," *Ex Libris* 222. marylaine.com/exlibris

Block devotes this column to her methodology in reviewing websites for *The CyberSkeptic's Guide to Internet Research.* As she points out, while most people would never do such extensive analysis, you *should* study your library's own website at this level of detail.

Without attempting to summarize an already-brief column, I'll note the "central questions" that Block tries to answer after studying a site:

> 1) If I was the person this site was trying to serve, would I find what I need and what I would reasonably expect the site to provide; and
>
> 2) Would I find it easily? Is the navigation intuitive and transparent, and is the search engine(s) up to the task?

Recommended.

Brewer, Joseph M. *et al*, "Libraries dealing with the future now," and Landesman, Margaret M., "Libraries investing in the future first—some practical suggestions," *ARL Bimonthly Report* 234 (June 2004).

The first (and longer) of these two pieces discusses a September 2003 retreat in Tucson to address the asserted need to "transform" academic libraries rather than "muddling through." I'm always uneasy with "transform," and recall the brief life of the ejournal with library transformation as its theme, but the issues raised here certainly deserve thought. The baker's dozen who took part in this exercise began with a dozen assumptions about institutions of higher education. Here are three of the twelve assumptions:

> 1. Institutions of higher education will experience a significant, long-term loss of budget and purchasing power over the foreseeable future.

8. Transformation will be "messy."

9. For change to occur, faculty and staff must perceive the likely future pain of an untransformed institution to be greater than the pain associated with making the transformation.

The first assumption is the key to this whole exercise: If it's false, there's no driving need for transformation. The other two listed here seem both certain and crucial.

As for the strategies suggested as relevant for each of three library types (muddling through, transitioning, and transformed), readers are better qualified to judge them than I am. I'm not aware that buying materials "just in case" is a wholly discredited practice, at least for university libraries with hopes of maintaining their long-term significance as something other than article-pushers. I'm not entirely convinced that "Believing digital is 'just another format'" is such a terrible thing. While I'd agree that transitioning properly involves changing what libraries count and measure, changing "what we value" seems a bit overreaching, unless that's at a very low level. "Creating a national network of regional repositories and libraries of record for print" says to me that books (and print in general) are regarded by this group as an annoying necessity, to be marginalized as much as possible, but maybe I'm overinterpreting.

Or maybe not: Here's assumption #8 about the transformed or transitioning library:

8. Libraries will support hybrid format environments for some time, but in new materials there will be a continuing shift to digital from paper and other tangible formats. *Libraries spend as little money as possible on adding to print collections.* [Emphasis added.]

I see nothing in the report that justifies this "don't buy print if you can avoid it" slant. Later, the piece enthuses that shoving all that print into "analog repositories" will free library space "for the creation of collaborative learning environments, shared faculty development areas…" and other spaces that seem to be part of good classroom buildings or "information commons." Maybe that is the future of the transformed academic library: Classrooms with reference librarians. It seems like a sad future to me, but I may misunderstand. **Read this yourself** and see whether you find it either agreeable or, as we are told in other words, inevitable ("libraries will not have a choice").

Margaret Landesman's accompanying piece offers sharply worded and highly useful advice, whether you buy into the "who needs physical collections in academic libraries?" future or not. For example, she suggests that librarians think more clearly about what things cost—more specifically, that increases in journal subscriptions be considered in *dollar* terms, not just percentages. If a $1,000 journal increases its price by 5% and a $300 journal increases its price by 8.3%, the $1,000 journal is taking twice as much more out of the library's funds--$50 rather than $25. As Landesman points out, we don't have the same problem when it comes to personal finances: "Tickets to the local opera cost just what they've cost for some years. Movie tickets have gone up substantially. I am not, however, tempted to believe that the fact that the opera did not increase its price makes it the more fiscally conservative choice."

She goes on: "The problem is the price, not the price increase." The *real* question with the $1,000 journal now offered for $1,050 is whether $1,050 represents good value—or whether that money might better be spent on other journals (or books?), including those from smaller publishers that have kept prices low but show higher *percentage* increases. The next point argues for libraries to find ways to turn ongoing costs (access and subscriptions) into "(mostly) one-time costs" (such as book purchases), e.g., by endowing certain digital resources such as the *Stanford Encyclopedia of Philosophy* to assure long-term access.

That's just the beginning. Landesman questions whether Big Deals have increased user satisfaction as much as they've increased journal counts; cautions against punishing the majority of publishers (that produce reasonably priced books and serials) because of the behavior of "Elseviley Verlag," as one librarian names the key problems. She recommends library investment in transformative initiatives such as PLoS and BioMed Central (and here I agree that some degree of transformation is desirable, albeit not inevitable); notes that canceling print isn't always a good idea; suggests meeting user needs by means other than ever-increasing subscription expenditures; and recommends that libraries work to establish institutional repositories.

I may not be wild about the idea that Google is the most appropriate "new front-end to our collections—it's free, it works, it's all anybody uses anyway," and I'd like to see *some* mention that academic libraries collect more than journals, but overall this is **recommended** and offers excellent food for thought.

Edwards, Eli, "Ephemeral to enduring: The Internet Archive and its role in preserving digital media," *Information Technology and Libraries* 23:1 (March 2004): 3-8.

Is it reasonable to think of the Internet Archive (IA) as a true digital archive? It's ambitious—no question about that—and founder Brewster Kahle has good intentions. This article considers IA in

some detail, notes the extent to which IA fails to meet library criteria for archival control, and looks at some library-based digital archiving projects. Edwards says that Kahle "regards electronic disseminators of information as digital librarians," an unfortunate dumbing down of the term, and that he's even suggested a "code of ethics for digital librarianship." There's nothing wrong with the code as excerpted, but I'm not ready to label everyone who disseminates information as a librarian. Given the limited searchability of IA (only by URL, last time I checked: it's *strictly* a known-item repository) and IA's automatic removal of any page at the request of that page's author, it's hard to say just what IA really is—or what it will be for the long term (assuming that it survives for an archival period, that is, a century or more).

IA has taken on a number of worthwhile initiatives. Edwards' article is **well worth reading** in its attempt to place those initiatives in a library context.

Moody, Kim, "Online portfolios, or "WOW! Look at everything I've done," LIScareer.com, June 2004. www.liscareer.com/moody_onlineportfolios.htm

> Do you have trouble recalling the many and varied experiences you've had as a library student or library professional? Feel like you're learning things at a rate of knots, but when you're actually asked, in a job application, to demonstrate your skills, you can't think of anything concrete to write? One solution is to create an online professional portfolio.

That's the lead paragraph in a seven-page article that goes on to describe a portfolio in areas beyond the visual arts, show its benefits, consider why it should be online, discuss what to include and what *not* to include, and finishes with a list of nine good pointers for building an online portfolio. The latter range from "Less is more" (the virtues of simplicity in website design) to "Spelling and punctuation are still important in cyberspace!"

It's a little late for me (and a portfolio that included "any articles or papers you have had published" would be ridiculous), but it's an intriguing idea for those earlier in their careers—and even for people well along who have never gathered together a record of accomplishments.

I could argue with one or two items, but those are mostly the mild hyperbole one might expect from a new Australian library school graduate. I don't think doing a portfolio online is likely to save "acres of trees," and I think it even more unlikely that sending your portfolio URL to potential employers will avoid "overloading them with half a rainforest worth of paper"—particularly since employers are likely to print out papers they find compelling. Extreme turns of phrase aside, this is a thoughtful (and highly readable) piece on an idea I think makes sense for most newer and potential librarians. **Recommended.**

Tang, Jinshan, Sridhar R. Avula, and Scott T. Acton, "DIRECT: A decentralized image retrieval system for the National STEM Digital Library," *Information Technology and Libraries* 23:1 (March 2004): 9-15.

"STEM" stands for "Science, Technology, Engineering, and Mathematics Education" and the name of this NSF-funded digital library can be further abbreviated to NSDL. The digital library aims to support education n the sciences and is "expected" to include tens of millions of images in its decentralized form. This paper discusses a "content-based image retrieval (CBIR)" system—one that claims to retrieve images *based on the images*, rather than the text in accompanying metadata.

True image retrieval is one of the dreams of information science; many systems have been developed, particularly over the past decade. How successful are such systems? That's always been interesting and never been clear. Those of us who don't believe censorware can be truly effective, particularly for CIPA requirements (which only involve images), question whether it's possible to recognize what an image is by working with only the image—particularly when distinctions need to be made between, say, classic art involving nude or semi-nude people and contemporary pornography of a type deemed harmful to children.

I don't think DIRECT gets into such details. Its "feature-based image retrieval" is based on color histograms and texture measurements. The paper goes into some of the details on how "features" are extracted and evaluated for retrieval. Examples suggest that, given one photo including a body of water, DIRECT can successfully retrieve other images representing bodies of water. Could it do equally well at, for example, locating other bridges given one bridge within a picture? Only long-term experimentation will tell. There's not enough here to claim that the image-retrieval problem has been solved, but it's an interesting look at one ongoing attempt.

Understanding Metadata, NISO Press, 2004. ISBN 1-880124-62-9. 16 p. Downloadable from www.niso.org; available as a hardcopy booklet from NISO Press.

Bad points first: The formatting is dreary—three tight columns of justified sans serif type (Helvetica or something equally boring), with every paragraph

indented even when it follows a heading. But the content is excellent: An introduction to metadata that should provide just enough background to get you going. Here's the first paragraph, a reasonably concise definition for those who find "data about data" insufficient:

> Metadata is structured information that describes, explains, locates, or otherwise makes it easier to retrieve, use, or manage an information resource. Metadata is often called data about data or information about information.

The second paragraph includes a key point for traditionalists: "Traditional library cataloging is a form of metadata; MARC 21 and the rule sets used with it, such as AACR2, are metadata standards."

The second section discusses what metadata does: Resource discovery, organizing electronic resources, interoperability, digital identification, archiving and preservation. After a brief discussion of structuring metadata comes the longest section: Metadata schemes and element sets. This includes reasonably detailed descriptions of Dublin Core, TEI, METS, MODS, EAD, LOM, <indecs>, ONIX, CDWA, VRA, MPEG-7 and MPEG-21, and metadata for datasets. **Read this booklet** and you'll know what all those acronyms mean—and may have some sense of why there are so many.

There's more: A discussion of creating metadata and quality control issues, more on interoperability, including crosswalks and registries, and notes on future directions. Get the booklet from NISO or download and print it; it's worth keeping around for future reference.

Library Stuff Perspective

Information Commons?

Kranich, Nancy, *The information commons: A public policy report*, Free Expression Policy Project, 2004, 58 pg. (Downloadable from www.fepproject.org.)

> The Internet offers unprecedented possibilities for human creativity, global communication, and access to information. Yet digital technology also invites new forms of information enclosure. In the last decade, mass media companies have developed methods of control that undermine the public's traditional rights to use, share, and reproduce information and ideas. These technologies, combined with dramatic consolidation in the media industry and new laws that increase its control over intellectual products, threaten to undermine the political discourse, free speech, and creativity needed for a healthy democracy.
>
> In response to the crisis, librarians, cyber-activists, and other public interest advocates have sought ways to expand access to the wealth of resources that the Internet promises, and have begun to build online communities, or "commons," for producing and sharing information, creative works, and democratic discussion. This report documents the information commons movement, explains its importance, and outlines the theories and "best practices" that have developed to assist its growth.

Those two paragraphs begin the executive summary for this 35-page report (the other 23 pages include resource lists, endnotes, and an index). A little later Kranich says, "Building the information commons is essential to 21st century democracy, but it is neither easy nor costless." Still later, the introduction says "large portions of the Internet were soon dominated by media corporations that developed 'technology protection measures,' licensing terms, and other 'digital rights management' techniques to restrict access to information and control its use. As a result, much online content is now wrapped, packaged, and restricted—treated as private rather than common property." That follows a comment about "dreams of a utopia where people could connect with myriad ideas and with each other instantly, no longer constrained by location, format, cost, time of day, on-site rules, or other barriers."

Right there, on the second page of the report, my red pen came out. While the internet isn't free—*someone* has to pay for all that infrastructure—I believe people *can* "connect with myriad ideas and with each other, no longer constrained by location, format…time of day…or other barriers." Millions of people use IM. Millions more use email, lists, weblogs, wikis—a variety of ways to communicate with each other and to set forth their ideas. I'm not sure "utopia" is the right word, but that particular set of possibilities is here, right now. Otherwise, I wouldn't know about the Free Expression Policy Project. Otherwise, *Cites & Insights* wouldn't exist. Does the internet *also* serve as a conduit for digital resources that carry direct prices? Sure it does. Among other things, that's how I make my living, but I never thought of RLG as a "media corporation," and I don't think I'd call OCLC, EBSCO or CSA media corporations either. I'm not sure how the existence of licensed resources on the internet threatens the future of open discourse on the internet.

I read the report twice, carefully, deliberately not marking anything the first time through. When I mentioned the "information commons" in February 2004, I said this:

> To date, it's not a concept that serves *my* mental models to draw other concepts together. I also wasn't terribly clear on a suitable definition. One definition was offered in [a commons-blog entry]; Mary Minow offered another definition; the discussion continues. I attended part of an ALA Midwinter forum on the information commons—and the por-

> tion I attended suggested to me that the concept continues to be ill defined.
>
> I'm sure this discussion will continue. Will I become an advocate for the information commons? Not directly, not until the mental model makes sense to me—but that could change at any time.

What better time to reconsider my position than now, with the issuance of this public policy report? I was hoping that Nancy Kranich would convince me that "information commons" was a well-defined concept and one that I should support.

That didn't happen—and I'm not sure whether it's because I'm unable to recognize the grand vision or because I don't buy this particular aggregation of concepts. The best I can do here is offer a few notes and queries and **recommend** that—if you think this concept *might* have merit—you acquire the report for yourself and make up your own mind. Sometimes it takes me a long time to "get it." Sometimes I never do. That doesn't mean you shouldn't.

The report has three major sections: Opportunities and challenges of the information age, the emerging information commons, and the future of the information commons. The second section is by far the longest, most of it examples of "open democratic information resources." I see a variety of different "commons" discussed, but I don't see enough shared characteristics to make it useful to think of them as all part of one "information commons."

Examples discussed include "software commons" or open source software; "licensing commons" (the GNU General Public License and Creative Commons licenses); open access publishing; OAI repositories; "institutional commons"—well, you'll just have to read that section and see what the examples have in common; and "subject matter information commons" or topical digital resource projects.

I look at that set, read the descriptions, and see an indigestible hodgepodge: A diverse array of initiatives (or in some cases independent projects that can be grouped into apparent "commons") that—to my mind—have little to do with one another except that they're not entirely for-profit initiatives. I try to connect the set of examples to the earlier discussion of the history of the commons and why "commons analysis" is important, and I can't make the connection. Again, maybe you can.

I support open access publishing (with some reservations) and OAI repositories (with greater reservations), at least to the extent that either or both might improve library budgetary situations and access to scholarly resources. I use a Creative Commons license and have written way too much about the need for more balanced copyright, including ways of adding to the public domain. I don't see those two areas as tightly coupled, and I *really* don't see any advantage to adding open source software to that mix. Stated as one big mass, the "information commons" is so huge and amorphous that it's hard to understand, and I think even harder to lobby for or make progress toward. Lobbying for the Public Domain Enhancement Act should proceed based on the many values enhanced by adding to the public domain—not, in my opinion, by focusing on the "information commons."

Some individual items bothered me. There's a claim that "the way the Web's portals and search engines are constructed may actually exacerbate, rather than remedy, the effects of media concentration by making it tougher to find all those independently created resources that are now available online." I haven't read the study that makes that claim, but I find it improbable. Between online directories (including the Librarians' Index to the Internet and Open Directory Project and), topical directories to lists, user groups and weblogs, and the very high visibility of weblogs and groups within Google, it's never been easier to find freely-available online resources—even though it's now also easy to get swamped by them. In fact, search engines *favor* freely available online resources; major licensed resources are almost always part of the Invisible Web, not available to web search engines.

Here's a paragraph that left me befuddled; maybe you'll see the point(s):

> The Internet facilitated not only expression "as diverse as human thought," but "peer production"—that is, decentralized production and distribution of information that bypasses the centralized control of more traditional publishing. As the legal scholar Yochai Benkler writes, peer production is "a process by which many individuals, whose actions are coordinated neither by managers nor by price signals in the market, contribute to a joint effort that effectively produces a unit of information or culture." The result is commons-based production of knowledge that, while not challenging individual authorship, fundamentally alters the current system in which commercial producers and passive consumers are the primary players.

I don't see the final sentence as following from the rest. As with most new internet resources, "peer production" adds "knowledge" that might or might not compete with traditional media, but will only fundamentally alter the producer/consumer landscape if *most people* find those resources more valuable than traditional media. An earlier paragraph says ten corporations control most of America's traditional media outlets. But the totality of those outlets—all the magazine titles, all the radio and tv stations, all the daily newspapers—comes to considerably less than a million (excluding magazine titles, the number would be under 100,000). There are *millions* of weblogs today. If citizens *preferred* those weblogs to tra-

ditional media, then the current system would indeed be "fundamentally altered." But if we *choose* to be "passive consumers" (a somewhat snide label for book and newspaper readers, but there it is), then no fundamental change will occur.

Nothing that Big Media has done has made it more difficult to find weblogs, participate in lists, join online groups, collaborate on wikis or write internet-distributed zines. But I'm guessing that the 100th best-read weblog, which *could* reach an audience of more than a billion all around the world, has a *lot* fewer readers than the 100th largest-circulation magazine or newspaper. I can't blame Big Media for that, much as I'd like to. Peer production needs to compete by offering more interesting and engaging resources than Big Media; all the "commons" in the world won't do that job.

I would be uncomfortable handing this pamphlet to a sympathetic Senator or Congressman who is also a careful reader. It's appropriate to argue that "intellectual property" is not identical to physical property and should not be accorded the same protections. But on page 9, Kranich quotes legal scholar Carol Rose who "counters that *property* regimes and even *individual property holdings* are 'by no means self-evident constructs'; instead, they are 'property arrangements that people have quite consciously talked themselves into.'" That reads as an attack on private property itself, which is likely to leave most policymakers even colder than it leaves me.

"Libraries are quintessential examples of institutional information commons." That's a pull quote in red type and it bothers me. Libraries may be *resource* commons, but libraries do *and must* deal in far more than information. Calling a library an "information commons" seriously devalues the library.

Is there one "information commons" or many semi-related "commons"? The recommendations and strategies that close the text of this report talk about "*the* information commons." Which presumably includes open source software and all the rest.

I probably agree with more than half of the recommendations and strategies—but as applied to individual problems, not as sweeping generalizations. I surely agree that noncommercial resources and new means of communication should flourish and become more meaningful over time. But the whole "movement" doesn't persuade me.

I would note in passing that "information commons" has at least two different meanings within the library community, as Rory Litwin pointed out in a *Library Juice* essay (July 6, 2004). I don't necessarily agree with the thrust of that essay, but it's important to note that "information commons" is commonly used to refer to "an area of a library with many computers for the public to use to access the internet and work with a variety of software." Substitute "students" for "the public" in the many information commons on campuses.

To try to make more sense of the "information commons" concept, I read an article by David Bollier, one of the leading lights of "information commons" thinking (cited in Kranich's report), "Why we must talk about the information commons." (*Law Library Journal* 96:2, 2004, readily available online). While I could add some questioning comments on that paper, I won't; I'll just say that it didn't help convince me.

The Good Stuff

Into the blogosphere: Rhetoric, community, and culture of weblogs. blog.lib.umn.edu/blogosphere/

If you search "Into the Blogosphere" as a phrase in a major web search engine, you'll find hundreds or low thousands of references, mostly pages that talk about being thrust "into the blogosphere" but some pages that refer to this collection of scholarly papers about weblogs, which came out right around July 1, 2004.

What you *won't* find—or at least I didn't within the first 300+ entries at the end of July 2004—is much in the way of critical discussion or review by anyone who's read the collection. I'm afraid you're not going to find that here either. I've read or at least skimmed all 20 papers and I originally planned to offer commentary on some of the papers or at least on the collection.

Remember text-e? My commentary on that trilingual colloquy took the equivalent of one entire issue spread out over three issues during 2002—and when I got to the final part, I realized that I should never have begun the commentary. This time, I had the good sense to look through the whole set *before* committing to any thoughtful analysis. I think it's beyond me to provide such commentary.

Into the blogosphere is an edited collective scholarly "book" published *as a weblog*, with comments allowed on each refereed paper. The length of some papers belies the notion expressed in one or two of them that weblogs consist of brief entries, but never mind. There's a foreword and an introduction; after that, papers appear alphabetically by the first author's name. That's probably more sensible than topical clustering, given the overlaps and oddities encountered here.

Some papers are scholarly in a way that reminds me that, although my college degree is in rhetoric (a

key element of the scholarship here), that was a *long* time ago. I'm no scholar these days, and if the more scholarly papers represent the scholarship of rhetoric, it's unlikely I'll ever return. Other papers are lighter on the academese and closer to ordinary English. I found some papers interesting and persuasive in their conjectures and analysis; others—well, who am I to judge?

One suggestion for potential readers, at least those who happen to be white males and fail to be deeply ashamed of that fact at every waking moment: *Skip the foreword*, at least initially. If I hadn't printed out all of the papers before reading the foreword, I would have run from the collection in horror. Maybe quoting the first two sentences will give you a sense of whether this is your cup of tea:

> Blogging offers one powerful way to embed a reraced, regendered liberal arts. The familiar system of studying/performing/credentialing is, *as folks reading this piece know*, premised on the magic number seven. [Emphasis added.]

I'm clearly not a "folk" who should be reading that piece and I suppose it's helpful in saying "only *our kind* should be reading this at all." I didn't know that Macedonians of ancient Greece were not white ("…embedded ideologues such as Aristotle, who was not, it should go without saying, white…"). I didn't realize that racism was "so deeply rooted in...the structures of the electricity, hardware, software…"—somehow, electricity never struck me as racist. I wasn't even aware that mathematics and astronomy were liberal arts. So I'll leave you with the conclusion of the foreword:

> With the 4 E's (explain, enable, embed, and enthymeme the verb) and the 7 reraced and regendered liberal arts (frequently presented as general education programs), as well as with the many suggestions, theories, insights, and inquiries of volumes such as *Into the Blogosphere*, we might have hope.

At that point, I lost all hope of making headway into the "volume" itself (an odd word for such an explicitly online collection). But I found several of the papers **well worth reading**. Make up your own mind.

Keizer, Gregg, "Busting the biggest PC myths," and Steers, Kirk, "Complete PC preventive maintenance guide," *PC World* 22:8 (August 2004), 107-14 and 152-4.

Keizer's article is fun and includes a cute "Bogus-O-Meter" for each of 15 "myths," but it helps to take some of it with a grain of salt. Specifically, although "Using a cell phone on a plane interferes with the navigation and communications system of the aircraft" gets a 4 (of 5) on the Bogus-O-Meter (where 5 is totally bogus), the article cites *evidence* that simulated cell phones can interfere with aircraft navigation and communications systems. Sure, they quote a "veteran pilot who didn't want his identity revealed" saying, "From everything I've read, cell phones and most avionics shouldn't conflict" (note that qualifier *most)*, but I've read comments by *named* airline pilots who have experienced similar interference. If it's my butt in the seat at 37,000 feet, I'm going to trust the FAA a *lot* more than an anonymous pilot and a PC writer.

On the other hand, the article is probably right in saying that household magnets won't destroy data on any modern storage device (except diskettes). Static RAM cards (SD, CompactFlash, etc. are immune, as are writable optical media—and as for hard disks, Bill Rudock with Seagate notes, "In every disk there's one heck of a magnet that swings the head." I also agree with the "very bogus" rating for "Turning off your PC daily to save power shortens its life" and "The government reads everyone's e-mail." Overall, a fun and perhaps valuable read.

I mention the Steers piece because of an odd expert-vs.-expert battle. Most of the tips are reasonably good, although a lot of people would disagree with the desirability of frequent defragmenting for contemporary (XP/NFS) hard disks. Here's the oddity, the second and by far the longest note in a sidebar "Four tips for longer PC life": "Leave your PC running." Steers subscribes to the hoary assertion that "Powering up from a cold state is one of the most stressful things you can do to your system's components." Steers' evidence? "I find that my PCs last longer when I keep them in hibernation." There we have it: From one, the world.

Manafy, Michelle, "Hey, pass it on!" *EContent* 27:5 (May 2004): 5.

Here's a concept: Treat "pass-along" usage as a sales opportunity instead of bending heaven and earth to make such "piracy" illegal and impossible. According to this editorial, some software publishers "seem to have gotten the 'customer isn't the enemy' message loud and clear." They're reinterpreting pass-along as a sales opportunity:

> Say someone dupes a software disc for me to try a program. When I enter their serial number and password, instead of being able to install and use the program or being informed that the product is registered to another user and that I'm forbidden to use it, I'm welcomed and cordially invited to demo the product free of charge for 30 days. Thus, my interest in the software is recognized (given my willingness to skirt legality by trying out my friend's software), and, without any additional steps, I can install a demo version that will soon expire and prompt me to buy.

Manafy goes on to say, "I respect their flexibility and perception of any user as a potential buyer. Distribu-

tion dynamics have changed and so must the digital content sales and protection proposition."

This ends an editorial that discusses Manafy's own habits with media: Relying on other individuals to spot things she'd be interested in, borrowing it if possible, and—if she likes it—buying other stuff by the same artist. She believes (correctly, in my opinion) that this behavior is not uncommon and notes that magazine publishers *rely* on pass-along readership as part of advertising rate calculations.

It's not clear whether the software DRM model noted would work for movies or music—but it *is* clear that millions of us are fed up with being treated like thieves by the RIAA and MPAA. There must be a better way; this one-pager suggests some possibilities.

Peer review and the acceptance of new scientific ideas, Sense About Science, 2004, 62pp. ISBN 0-9547974-0-X.

This "discussion paper from a Working Party on equipping the public with an understanding of peer review" is first rate and **well worth reading**, even if it seems a tad long (my copy comes in at 41 pages; I must have discarded some appendices). You can download it (PDF) from www.senseaboutscience.org.

Sense About Science is a relatively young UK "registered charity" (what we'd call a nonprofit in the U.S., I think); one of its objectives is to promote public knowledge of how scientific research is conducted. The working party, chaired by Prof. Sir Brian Heap and including Dr. Derek Bell, Ms. Tracey Brown, Prof. Stevan Harnad, and nine others was convened in November 2002; this report appeared June 24, 2004.

I read it in stages and found myself marking more than fifty paragraphs—not to criticize them, but to point out particularly interesting segments. The wide-ranging, well-written report discusses the dangers of "scientific" claims in the public eye that have *not* been peer reviewed, admits that peer review doesn't really guarantee quality (given the hierarchy of journals down to "virtual vanity press," almost any paper can get published *somewhere*) or prevent fraud, and explains in considerable detail how and why peer review does and should work.

Section 2, "A guide to peer review and scientific publishing," is sixteen pages and should be read by anyone who doesn't understand what it's all about. It's still on the long side; a shorter version will be available in October 2004. Meanwhile, anyone with a serious interest in the STM literature should take a look at this full document. The price is right, it's nicely done, and you're almost certain to find tidbits here to use on those who regard peer review as some elitist notion or relic of the "Gutenberg era." **Strongly recommended**.

Perez, Ernest, "High-powered note-taking with hand-held pen scanners," *Online* 28:4 (July/August 2004): 27-30.

This one just plain surprised me. I remember early handheld scanners with OCR support and how amazing it was when they provided any useful recognition. Technology marches on. These "pen scanners" are cheap and appear to provide good results in appropriate circumstances. As usual with Perez, the article is well written and free of excess hype. **Worth reading** as a suggestion of a technology you've probably never thought of, that just might be ideal for some projects.

Srodin, Sharon, "Let's make a deal!" *Online* 28:4 (July/August 2004): 16-19.

The subtitle is "Tips and tricks for negotiating content purchases," and it's an interesting piece. Fortunately, "contract terms in the information industry are usually nondisclosable" *doesn't work* for public institutions (including universities and colleges) in a number of major states, including California, so there's some hope that a database of real-world prices could happen. In the meantime, Srodin offers seven tactics to improve negotiation, from mooching off of someone else (if you're in an industry where there's likely to be more than one "procurer of third-party content") to name-dropping, emphasizing how wonderful your outfit is and how privileged the third party should be to deal with you. Some of the tips may not apply, but it's certainly **worth reading**.

The Details

Cites & Insights: Crawford at Large, Volume 4, Number 11, Whole Issue 54, ISSN 1534-0937, is written and produced by Walt Crawford, a senior analyst at RLG. Opinions herein do not reflect those of RLG. Comments should be sent to wcc@notes.rlg.org.

URL: cites.boisestate.edu/civ4i11.pdf