# Cites & Insights

## Crawford at Large
## Libraries • Policy • Technology • Media

## Ethical Perspectives
# Republishing and Blogging

Let's talk a little about ethics in two different spheres: journal publishing and blogging. This may be the start of a new topical theme; it may not. If you're one of those tuned in enough to assume that "ethics and blogging" means a long analysis of the "Webcred" closed conference—well, that was my original intention, but no such analysis appears here.

### The Ethics of Republishing

Easily the most astonishing presentation at last November's Charleston Conference was Philip M. Davis' preview of this paper. (Full title: "The ethics of republishing: A case study of Emerald/MCB University Press journals," to appear in *Library Resources & Technical Services* 49:2 later this year.)

Davis had anecdotal evidence that Emerald (formerly MCB UP) had republished some articles in its many journals without appropriate notices. He did some keyword searching within Emerald's online journals—and identified "409 examples of articles from sixty-seven journals that were republished without explicit identification from 1989 through 2003."

There's nothing inherently unethical about republishing an article. Freelance writers reuse their own material all the time; that's essential if they want to eke out a living. Refereed journal articles are a different story, but it's still reasonable to republish—providing the republishing is transparent and meets certain ethical standards.

Davis quotes Joseph Fulda's suggested guidelines for ethical multiple publications. Briefly, the five guidelines are that republication should be in journals representing different fields; the editor of the second journal knows the article has been previously published; prior publication is explicitly acknowledged in the second publication; the duplication isn't simultaneous; and the journals don't have overlapping readership. There are other guidelines, but all have in common that the republished article must be explicitly identified as such.

## Inside This Issue

Fulda's guidelines fail in certain cases. Landmark articles will be republished in the *same* field and possibly the same journal simply because they are landmark articles. The *Information Technology and Libraries* issue celebrating LITA's 25th anniversary republished a few "classic" articles from previous issues—explicitly identified, to be sure. (According to Davis' article, most "redundant publication" within the medical literature is *not* exact duplication—which makes redundant publication even worse, because it's much more difficult to identify.) Davis offers examples of legitimate republishing—and notes that he found nothing in the literature about publishers duplicating articles within their own journals without explicit notice.

How bad was the Emerald problem? "Most articles were discovered to be published in two journals simultaneously…or after a significant delay." In some cases, an article was republished in the *same* journal without explicit labeling. Some republished articles had slightly modified titles. Articles were duplicated

across journals within the same general field and across journals in closely related fields.

Some instances seem bizarre. *Library Management* 16:5 and *Management Decision* 33:5, both 1995, consisted of the *same ten articles*—and these are by no means inexpensive journals. Another 1990 case had two different journals with identical articles for one issue. *All* of the articles in *Asian Libraries* 6 (1997) had appeared in other library publications. Twenty-seven of the 40 articles published in *OCLC Systems & Services* in 1997, 1998, and 1999 were duplicated in other journals. Those aren't the only cases.

Do the journals reach entirely separate audiences? Well, the subscription lists certainly overlap (based on RLG Union Catalog holdings), as you'd expect.

Did authors, editors, and editorial boards know this was going on? Some authors did; some claim that they did not. Some editors didn't respond, some were new to the journals, and one provided a labored explanation for duplicate publication. Contacted librarians who'd served on the *OCLC Systems & Services* editorial board weren't aware of the duplication.

As Davis notes, this isn't about illegality—but unacknowledged duplication harms the subscriber, particularly within expensive journals. When you subscribe to journals, you're paying in advance for what you normally assume will be original content.

## Emerald's response

Davis offered Emerald a chance to respond during the Charleston Conference, which they did; they also posted a written response. (That response indicates that Emerald "extended our full cooperation"—although Davis' article suggests that "full" might be an exaggeration, based on unanswered questions.)

The response agrees that explicit notice of duplication should have been published "and regret any inconvenience as a result of that notice not being given." It says authors *were* informed, and that it was practice between 1989 and 2000 to republish articles "within another MCB journal where it was felt that their content would be of interest or benefit to the additional journal audience." That doesn't explain republication within the *same* journal, but that's admittedly a small portion of the whole.

The response also excuses republication as a way to "help" where newly-acquired journals had a "significant delay in the despatch schedule." That's tricky; one would not normally think that buying a journal

gives a publisher license to temporarily fill its issues with already-published articles.

Emerald says there has been no deliberate dual publication since 2001 and that the corporate policy is not to practice dual publication (except in special cases such as anniversaries and landmark papers).

## But wait, there's more

That might be the end of it, particularly since Emerald renamed itself; the live rejoinder had a feel of "the bad old days at MCB UP," separating the speaker from those problems. Or maybe not.

Davis prepared a followup letter, which should appear this summer in *Library Resources & Technical Services*. In that letter, he says there were *numerous* instances of republication within the same journal; that he's now identified duplication going back to 1975; and that he finds significant subscriber overlap among journals in which duplication appeared. The number of journals involved is now up to 73.

Most of the letter deals with a related problem that also raises ethical issues. It involves another publishing company that appears to have strong ownership overlap with Emerald (and is purchased through Emerald)—and the fact that owners were also functioning as editors *and authors*. In fact, Dr. John Peters republished his own article from one MCB UP journal in another journal that he edited at the time.

There's nothing inherently wrong with owners being editors and there's nothing wrong or unusual about editors being authors. In refereed journals, however, unless the refereeing process is known and unusually transparent, it's tricky when an editor's own work appears within the refereed section. Davis offers a slightly stronger statement:

> The peer review process—which is at the heart of scholarly communication—has been cast into doubt. Furthermore, conflicts of interest when individuals serve as owners, managers, editors and authors of academic journals lead us to question whether these individuals may not have been acting in the best interest of scholarly communication. Commercial interests have outweighed editorial independence.

## Full disclosure

I have no current involvement with Emerald. I published articles and columns in almost every issue of *Library Hi Tech* between 1984 and 1998, and served on its editorial board from 1986 through 2000. The predecessor to *Cites & Insights* appeared in *every* issue of *Library Hi Tech News* between March 1995 and De-

cember 2000. MCB University Press purchased the two publications in mid-1998. When MCB University Press substantially increased prices for both publications, I did not immediately leave the editorial board and terminate *Crawford's Corner*. In the end, I left the editorial board because I'd been on it long enough—and terminated *Crawford's Corner* over a personnel issue. I do *not* have any personal experience of ethically questionable behavior by MCB University Press during the time I was writing for them or serving on one of their editorial boards.

## Ethics and Blogging

Some library bloggers have been writing about the ethics of blogging—and a model Blogger's Code of Ethics appeared recently at CyberJournalist.net. We'll get to that. First, in chronological order (as usual), comments from Michael Stephens and Karen Schneider and some responses to their posts.

Last June, Stephens' *Tame the web* included one of his noteworthy "numbered lists," this time "Ten things a blogging librarian must do (an exercise in common sense)." Omitting the expansions, here's the list: Cite your sources. Post often but have something to say. Make the commitment to follow through (that is, keep your blog active). Post about what you're passionate about. "Share yourself." Show your administration how well an external blog is working. Don't do a personal blog on your library's dime. "Blog unto others as you'd have them blog unto you." Read other blogs for inspiration. Learn all there is to know about your software. And have fun." (OK, so it's 11.)

It's an interesting list. One commenter suggested that "not doing it on your library's dime" was unrealistic. Another emphasized the need for bloggers to *keep* blogging: "It's really disheartening to read some really nice first posts, quickly bookmark/blogroll it, only to later discover that no newer posts have been published since." The rest of the comments were spam.

Stephens returned to this topic on November 14, with a suggested list of "Library blogger's personal protocols"—six items, this time, with more comments on each. Respect your organization. "Don't be afraid to put your two [cents] in on something you *really* believe in." Play nice (cite sources, link back). Don't reveal secrets. "Blog anonymously… or blog proudly and let your administration know what you are doing." That last is paraphrased from two separate protocols: You need to choose one or the other.

I mentioned Karen Schneider's first couple of posts on "Blogging and ethics" in *Cites & Insights* 5:1. She continued on December 13 with Part 3, "the anti-guidelines." It's a 15-point set that may slightly overstate the things bloggers do wrong, but it's also funny and **worth reading** (you can find it in the archives at freerangelibrarian.com). She raises a bunch of ethical issues indirectly, from getting paid for promoting products (without telling readers about it) to avoiding accountability. It's a good list to avoid.

Karen got into deeper waters—fundamental ethical issues—with Part 4, "Don't stand too close to me." Say she's helping to pour punch and serve baked goods at a party for church parishioners and overhears interesting conversations—maybe picks up some great gossip. Is she free to blog about them? She doesn't believe so, but notes an article by Jeffrey Rosen offering examples of bloggers who feel perfectly free to invade other people's privacy. She also has a personal example, "when a librarian blogger quoted another librarian on an issue I was sure was confidential." (Karen causes no further harm—the item is so blind that I can't imagine any but the parties involved recognizing it, and probably not them.)

> We librarians are all about free speech. But the First Amendment won't make you less of a chump for kiss-and-tell blogging, and it won't expunge the stain on your professionalism for knowingly crossing the line between private and public.

So far, I haven't seen anyone disagree with what strikes me as a pretty fundamental ethical assertion.

Finally, there's "A blogger's code of ethics" at *CyberJournalist.net*: www.cyberjournalist.net/news/000215.php. This proposed code of ethics, based on the Society of Professional Journalists Code of Ethics, includes nineteen bullets in three general categories: Be honest and fair; minimize harm; be accountable. Specifics involve avoiding plagiarism, distortion, and misrepresentation, identifying sources when feasible, distinguishing between facts, advocacy, and commentary, and distinguishing all those from advertising. Ethical bloggers should show compassion for those who may be affected, avoid arrogance when pursuing information, recognize privacy as a general right, and show good taste—and they should admit and correct mistakes, explain a weblog's mission, disclose conflicts of interest, and "abide by the same high standards to which they hold others." It's a strong list, **well worth**

**reading and printing out** if you do a weblog. (Or a web-based print journal, for that matter!)

## Comment

Where does this all leave us? You'd be surprised how many ethical quandaries can be resolved by paying attention to the single law/rule/ethos that pops up in almost every religion and moral system: Treat others the way you'd wish them to treat you.

I do republish columns—*always* with explicit identification and some good reason to do so. I probably fail to clearly distinguish between facts, advocacy and commentary: It's tough in a journal such as this, where *everything* is partially commentary and advocacy sneaks in unexpectedly. I certainly don't claim ethical perfection; I do claim to care about ethics and to strive for ethical soundness. Most library weblogs that I read—even the ones I vehemently disagree with—do pretty well, most of the time. Maybe I've been lucky. Maybe most librarians really are professionals who understand what professionalism is all about. I believe the latter.

# The Library Stuff

Holt, Rachel, "Musings of a nextgen librarian," *LISjobs.com*, January 2005 (www.lisjobs.com/newsletter/archives/jan05rholt.htm)

Holt offers characteristics that "we"—apparently, *all* MLS holders in a certain age range—share. When has there been a generation all of whose members share all primary values? What are those key characteristics? "First, and most importantly, we have *chosen* to be librarians" (as opposed to taking it on as a second career). [Emphasis added.] "Second, and perhaps most obvious, we have a very high level of comfort with technology…. Third…we have a shared cultural experience with the generation following us, the Millenials…. Fourth, we are the first generation of librarians to have experienced the library in its most modern mutation. We were the first students to use online library catalogs… Fifth, and most enjoyably, we bring a strong desire to do away with dusty librarian stereotypes and revise the profession for the culture at large…" Before that list, Holt tells us that Nextgens are "often viewed by our colleagues with a mixture of bewilderment and mistrust… Then we open our mouths and say words like "social responsi-

bility," and boy, you can just hear the arms crossing and the noses sniffing in libraries everywhere."

I'm not sniffing (well, I am; I'm getting over a cold), but I am bewildered. Online catalogs have been around for more than two decades now, so either Nextgens are older than I think, or there's a sloppy claim there. I *know* that SRRT precedes Nextgens by quite a few years: "Social responsibility" (whatever its merits) is not a rallying cry new to this generation. There are a lot of older librarians with a "very high level of comfort with technology"—and I'll go out on a limb, but I'll bet there are some "Nexgen" MLS-holders who aren't all that comfortable with "computers as helpmeets, avatars, gateways, and tools."

What's bothersome in this screed is the implicit "*and you're not*" that runs throughout it. Holt asserts that people who *enter* the workplace as librarians are "unique in many of our institutions" and strongly implies that (all? most?) older librarians reached that path "through recession or industry belt-tightening." She certainly implies that older librarians just can't be as tech-savvy as Nextgens.

Holt and (all?) her peers "want to sit at the grown-ups' table." Part of that might be getting over your good self and how you and your peers deserve special treatment. If you're special, show it by doing good work. "Offer a fresh voice" by all means—but make that voice count by saying important things, not by harping on how special you are because you're young (and we're not). The antepenultimate sentence of this three-page article (which I **recommend that you read**, since I'm probably being way too harsh), reads: "We're not trying to steal anyone's job, wipe away the traditions of the profession, or show disrespect to the people who have done this before and who do it better than we ever could." I have a little trouble buying the last clause of that sentence, and that's a shame.

## A walking paper cluster

Aaron Schmidt at *walking paper* (www.walkingpaper .org) has provided a series of thoughtful, provocative blog entries over the last few months. Last month, I mentioned his "top ten things to stay tech currents" and "tech needs pyramid." He's kept it up—and, in the first instance here, provided an indirect response to my July 2004 (*C&I* 4:9, p. 14) comments on his May 17, 2004 posting "Once bitten" (which discussed the extent to which library ebooks were being pushed

by producers more than requested by readers—while books on MP3 appear to meet a real user demand). Here's my thought, "which could turn into a full-fledged essay or article":

> In the case of ebooks (and particularly dedicated ebook appliances), libraries were "getting out ahead" of patrons—demonstrably, since the number of consumers who purchased ebook readers for their own use is so small that nobody's ever offered an estimate. My guess is that it's almost always a bad idea for public libraries to try to be ahead of their users in adopting new media, particularly new circulating media. Instead, I believe, it makes sense to be a little behind: Ready at the point where a new medium serves more than the most privileged set of "haves" in the community. But that's still rough thinking, and far be it from me to criticize library actions.

"Serving two masters," posted December 12, 2004, considers the fact that "many public library users…are not terribly adept with technology" while other users "work with technology on a regular basis and have fairly high technology expectations." As he notes, there wouldn't be a problem "if we had unlimited fiscal and temporal resources"—but in the real world, libraries need to find balances. Schmidt notes, "There isn't one answer that will apply to all situations," since every community is different, but does offer four general suggestions: "Include technology in community surveys… Mine public services staff for information [on patrons' technology skills]… Do trial projects… Spy [see what's happening elsewhere in the community]." It's a **recommended** essay in general. I wish more people would mention the worth of true trial projects, "dip[ping] your proverbial toe in the water," with the recognition that a trial is *always* a success even if the end result is *not* retaining a new service or technology.

And here's Schmidt's indirect response to my off-hand thought last July:

> It is most realistic for libraries to aim to be as current as their surrounding community. Less realistic but perhaps more appealing is the notion of libraries being their community's technology mentor. While I think certain situations warrant that the majority of a library's attention be given to making sure people's basic needs are met, I think there are some scenarios in which libraries could lead their communities by purchasing certain types of technology. More on this later.

I hope Schmidt does a better job with his last sentence than I've done—offering more thoughts later—

because he's actually out *in* a public library and because I respect what he's saying. Not that there's a necessary contradiction: Libraries can be slightly behind the community in adopting new media while leading the community in certain types of technology. Maybe what I should have said is "slightly behind the early adopters." (At this point, I'm impressed enough with what he's saying that I'm considerably less likely to expand my own comment. Why do the work when someone else is doing it better?)

*Too much fun: December 16, 2004*
Here Schmidt comments on the "Digital Photo Effect," which he picks up from Rajat Paharia's *rootburn* log. Requoting:

> What's the DPE? My ability to produce and acquire has far outstripped my ability to consume…. This has a couple of ramifications:
>
> 1. I feel behind all the time.
>
> 2. Because there is so much to consume, I don't enjoy each individual photo as much as I did when they were physical prints. I click through fast.
>
> 3. Because of 1 and 2, sometimes I don't even bother.

Paharia notes that this is happening with music and video as well. I offered a mini-perspective last November, "Does the Music Matter?," riffing off a July 17 New York *Times* essay, "Can an MP3 glutton savor a tune?" In that essay, Roger Van Bakel notes that he doesn't make the mental connections with any of the thousands and thousands of MP3 songs he downloads that he used to make when he owned a lot less music. I suggested that Van Bakel might be on to something—that having so much of it so readily available may cheapen the emotional impact of the music.

Schmidt says DPE, a generalized version of MP3 overload/under-connection, "has come up in a number of conversations I've had with people over the last few months. I know a number of people [who] have more music than they know what to do with. They have only a vague idea of what is cached on their hard drives, and seem to be not too enthusiastic about most of it."

Conclusion? Maybe there is a real problem out there—maybe oversaturation does undermine deep connections. Schmidt goes on to bring this back to the library: Libraries can help prevent DPE in patrons through selective promotion, and the values that libraries add to simple media provision can "snap people out of their DPE slumber." I hope that's true. I

*know* public libraries add value and need to be aware of and promote that value.

### Rss hub-bub, January 19, 2005

This time Schmidt's just asking for trouble. Noting enthusiasm in the blogosphere about one library vendor adding RSS to one of their extended products (and the predictable "every library and every vendor should be doing this *right now*" responses from more excitable bloggers), he quotes part of one comment on one post. That comment, from an employee of another library automation company, notes that when that employee has suggested RSS feeds, the general response is "where are the customers who want this?"

> He has a point that is sometimes difficult to remember. There are still many, many people [who] aren't familiar with RSS. Ask your neighbor what "Really Simple Syndication" is. 98% of you will come back having received strange looks, and maybe 1% of you (likely less) will have the correct answer. [Footnote: The missing 1%? You'll come back with a black eye.]

You won't get RSS in online catalogs until vendors know that patrons are using it—and, by the way, you probably won't get it if you're not willing to pay for it. Sure, it has valuable library roles—but what portion of the community will take advantage of the feeds? Maybe, as Schmidt suggests, this is one of those cases where the library mentors the patrons—"guiding them through technologies they might benefit from learning about."

He also notes that, if RSS takes off in a big way, it's likely to be ruined—"If not by some new fangled spam, then it'll be by the abundant adverts and few full-content feeds. It could be rendered as painful to use as email." I've wondered about that, and noted with a small sense of irony that the RSS feed from one of the top library promoters of RSS feeds is now partially broken (by my standards): It's no longer a full-text feed, for financial reasons. (And, earlier, notes that he only encountered the comments because he clicked through to the site.)

Interesting stuff. So your library would just as soon drop its new title lists and substitute an automatically generated RSS feed? You tell your patrons, "Oh, we don't send that email any more. All you have to do is add our new title RSS feed to your aggregator." What reaction will you get?

I live in a *very* high-tech community, on a block where most homes are owned by two parents, both of whom work in Silicon Valley. If I went around asking neighbors about RSS, I'm sure I'd get more than 1% success rate—but I'm also sure it would be a lot less than half.

(Last-minute addition: See TRENDS & QUICK TAKES in this issue. The latest Pew Internet & American Life study on blogging suggests that Schmidt's "98%" figure is right on the money.)

### Some problems with virtual reference
### January 20, 2005

I don't know enough about how VR, chat reference, IM reference, email reference, and old-fashioned F2F reference actually work to comment actively in this area. I was a bit surprised by some earlier articles seeming to attack VR, particularly since one of them came from an extreme advocate of commercial VR services. In this post, Schmidt points to a more recent piece in *Library Journal* by Brenda Bailey-Hainer, "Virtual reference: Alive & well," which points to successes within statewide VR services—and argues strongly against abandoning other forms of reference in favor of VR.

The post considers some inherent problems with VR systems, and is **recommended reading** on its own. He notes that VR benefits librarians more than patrons; that the systems were clearly built from a library-centric perspective; and that statewide and nationwide VR systems do not connect patrons to their community. "Large scale projects can (somewhat) successfully answer patrons' general reference questions, but they cannot provide answers to local questions or handle home library specific tasks" (such as those having to do with holds).

Schmidt argues that libraries should try to find ways to be "where our patrons already are"—possibly including IM reference. He also agrees that librarians should find ways to make VR work better.

> Right now, though, it is just addressing a **symptom**—"We need to be answering questions for patrons online." Perhaps in the future it will be able to address the root **problem**—"We need to be **connecting** to our patrons online."

One of my predictions during the LITA Top Tech Trends discussion at Midwinter in Boston was that at least one or two significant new commentators would emerge from the library blogosphere this year. I gave Schmidt as one example, although he's been doing it for quite a few months now.

Webster, Peter, "Breaking down information silos: Integrating online information," *Online* 28:6 (November/December 2004): 30-34.

Webster provides a quick, clear overview of some "manifestations of an integrated information environment"—ejournals, open online catalogs, metasearch using Z39.50 and other protocols, metadata harvesting, OpenURL, and open access.

The section on "the new gray literature" is provocative and bears thinking about. Webster notes that much of what Yahoo! and Google find "would once have been gray literature." The controversial part: for many people even on university campuses, "anything not readily accessible via such a search engine has become the new gray literature, useful perhaps, but annoying and difficult to retrieve because it cannot be searched by standard methods." The scholarly corpus as gray literature: Disturbing, but I'm not ready to say he's wrong. **Worth reading and thinking about.**

## Trends & Quick Takes

# WiFi or Not WiFi?

Here are a couple of interesting and possibly related items (courtesy of Marydee Ojala at *Online*, who found the conjunction intriguing). A December 6, 2004 PR announcement from Airespace, Inc. announced that the Airespace Wireless Enterprise Platform has been deployed throughout the 79 libraries of Chicago Public Library "to deliver free Wireless LAN services to Chicago residents, library personnel, and mobile City workers."

Four days earlier, the Information Technology Association of America (ITAA) issued a press release criticizing a legislative deal to bring city-owned WiFi to Philadelphia residents. Why? Because the legislation included a deal giving network operators in the rest of Pennsylvania the power to block deployment of similar municipal WiFi networks. ITAA's President Harris N. Miller: "This is a clear case of an incumbent network operator using political muscle to limit new wireless broadband technologies. Backroom deals are not the way to resolve these issues—competition is."

I agree that the juxtaposition is interesting—and have no further comment.

## OPAC Manifesto?

Jessamyn's started something new at librarian.net: The OpacManifesto wiki. (www.librarian.net/opwiki). It's worth a look. I printed off the "manifesto" itself as of November 11, 2004, but I'm sure it's grown since then. The manifesto at that point included 11 bullet points under "what library staff wants," seven "what geek wants," and 14 "what users want." (For readers outside the library community: **O**nline **p**ublic **a**ccess **c**atalog, or just "online catalog.")

Some seem obvious: "We want our customizations to the OPAC to not be overwritten by the next OPAC release when we install it." "We [users] want the online version and the in-the-library version to look and act the same." Some seem a little tricky: "We want easy icons for distinguishing records. Ones that can be based on 008 or 300 or 6XXv or location in the holding record or…" Some relate directly to problems with too-clever designs: "We want the back button to work when we use the OPAC."

## Patent Holding Companies

A December 16 news.com story by John Borland notes that Acacia Research is buying Global Patent Holdings. So what? So this: Global Patent Holdings is one of those beloved companies whose only products appear to be litigation and licenses—companies that buy patents developed elsewhere, then make the broadest possible claims and threaten to sue any company deemed in violation of the patents.

As you should know, some technology-related patents are wildly overbroad—but for many companies, paying for a license is less expensive and less hassle than going to court and attempting to invalidate the patent. The story begins, "In the streaming media business, a letter from Acacia Research usually means one thing: the threat of a patent lawsuit." The purchase will make Acacia more of a "patent powerhouse"—the CEO explicitly says the goal is "becoming the leading technology licensing company."

Not "the company that creates the best technology and licenses it." Creation—"the progress of science and useful arts" as the Constitution calls it in the copyright-and-patent clause—isn't what these companies are all about. These companies produce licenses and litigation. (Former Microsoft CTO Nathan Myhrvold has founded a similar company, Intellectual Ventures, with close to a thousand patents already.)

I'm not wild about patent holding companies. Edward Felten disagrees, in a January 12, 2005 *Freedom to tinker* posting: "From a policy standpoint I don't see a problem." He makes some good points, *if* we're dealing with legitimate patents. Patent holding companies can provide a level ground for smaller inventors: True. Inventors should be able to focus on invention, not on extracting royalties: Also true.

As Felten says, "those who support rational patent policy should focus on setting up the right patent rules (whatever they are), and applying those rules to whoever happens to own each patent." He's right, of course: My outrage at patent holding companies is based on the *kind* of patents we hear about and the overbroad claims. If smaller companies and inventors actually do rely on patent holding companies to gain justifiable rewards for their real inventions, there's no reason to object.

The two comments I saw on the posting when I downloaded it (the day it was posted—there may be more since) both acknowledged this. Grant Gould noted what's needed to make the patent system "economically efficient" (and just from a policy perspective): "strong prior-art investigations, a more objective obviousness criterion tied to the likelihood of reinvention during the patent term, an independent reinvention defense to infringement claims, increasing renewal fees tied to the price of a license." "Skopo" says Felten "misses the point"—which is not that the holding companies have no other business but that some of the patents being enforced are overbroad. I don't know that Felten misses that point, but it's a good one. I withdraw my general outrage over companies whose only business is to enforce patents they purchase, although their aggressiveness may itself be a problem. The bigger problem is patents that are too broad and, in many IT-related cases, should never have been issued.

## Power for a Price

What do you get in a cheap name brand PC these days? *Computer Shopper* liked the $600 eMachines T3256 enough to rate it an Editors' Choice. (Yes, eMachines is a brand name; it's the same company as Gateway—but eMachines doesn't offer custom configuration.) That price, what you'd pay for a Motorola Razr smart phone, buys a 2.2GHz Athlon XP 3200+ (claimed to perform as fast as a 3.2GHz Pentium4), 512MB DDRAM, a 160GB hard disk, multiformat DVD burner and second CD-ROM drive, 64MB nVidia GeForce4 MX graphics (on the motherboard), Windows XP Home, MS Works, and basic speakers. Oh, and a card reader for your camera and other "media" memory. You still need a display, but that's a lot of computer for the money—and if you're a gamer, there's an AGP slot to plug in a hotter graphics card.

## Blogging and Triumphalism

Here's Pew again, once more extrapolating from 1,800 interviews to give us the precise state of the nation on internet-related issues. (Yes, 1,800 interviews chosen with appropriate tools *should* be enough for reasonably accurate projections, given a whole set of hard-to-test assumptions.) This time it's about the blogosphere. I didn't download or read the whole report, but I did look at the summary and some comments about the study and the summary. I'm assuming here that "adults" means "age 18 and over."

I'm going to repeat some of the key points in the summary, using *precisely* the information given, but wording them just a bit differently:

➢ 96% of U.S. adults have not created weblogs.
➢ 86% of them do not read (and, I would extrapolate, have never read) weblogs. 80% do not know what a blog is. 93% have never posted a comment or other material on blogs. During the political campaign, 95% of adults did not read political weblogs—and 97% did not read them regularly.
➢ 97-98% of U.S. adults do not use RSS aggregators or XML readers.
➢ 52% of blog creators are more than 29 years old.
➢ 58% of blog creators are not particularly well off financially, living in households with no more than $50,000 annual gross income
➢ 61% of blog creators do not have college degrees.

As some readers have figured out by now, I've just provided the inverse of the claims actually made in the summary—and adjusted for the difference between 120 million adult internet users and around 222 million adults (2000 census).

j Baumgart at *j's scratchpad* took issue with the way the figures were originally presented—particularly "blog creators are more likely to be" claims on education, where 39% is "more likely."

*The Register*'s Andrew Orlowski also took a swipe at the report, headlining the January 5 story "62pc of netizens unaware of Pajamahadeen militants" ("pahamahadeen" being one snarky word for bloggers, and I suspect particularly those bloggers who assert that blogging is making or has made traditional journalism obsolete). I don't usually read *The Register*; I picked up on this one because Orlowski credited Seth Finkelstein (sethf.com/infothought/) with "Finkelstein's Law," the ease with which a few people agreeing with each other gets multiplied into a Movement. You could also call it the echo-chamber effect. (I'm oversimplifying and possibly misrepresenting: go to archives/000516.html on *Infothought* for the proper interpretation.)

I mention this partly because of stuff I saw coming out of reports on the "Webcred" conference, partly because of a silly story by Adam L. Penenberg at *Wired News*, "Like it or not, blogs have legs." Well, of course they do. But does a medium ignored by 85% of adults *really* constitute a "revolution in the dissemination of intellectual capital"? Penenberg dismisses "solipsists…who once thought online news would never equal print." Well, then, call me a solipsist: While "never" is a strong word, I do not believe online news takes the place of good print newspapers or newsmagazines now. It does something different, which is as it should be.

Penenberg's on a roll here. He calls the blogosphere "at its best…a pure meritocracy." "In a sense, blogs function like peer-review journals do in the academic world." But better! Because blogs aren't controlled by a publishing cartel which "deleteriously affects the level and quality of discussion."

> With blogs, however, anybody with an internet connection can engage anybody else. Concepts are presented, attacked, sliced, diced, added to and subtracted from, mangled, massaged and molded until what is left is an amalgam of the finest we as an online society have to offer.

Whew. Who woulda thought? Where do you go to reap that refined amalgam?

Penenberg goes on to talk about Chris Anderson's "Long Tail" thesis, which is old news to book publishers but hot stuff when *Wired* publishes it. There's more; you can find the original without my snarkiness at www.wired.com/news/print/0,1294, 66336,00.html

Weblogs can be—no, *are*—important. Weblogs are useful. Their use and readership are growing rapidly. But they haven't overthrown any other medium—it doesn't usually work that way. And, for 86% of American adults, weblogs *don't exist*. (Interestingly, an offhand assertion by Aaron Schmidt, noted in THE LIBRARY STUFF, turns out to be almost exactly on the money: 98% of American adults don't know what RSS is, or at least haven't used it.)

## Quicker Takes

"There are two truisms in our go-go, tech-driven culture. First, every technological change improves our lives. And second, the faster the change, the better." Nope, not John Dvorak this time. Those are the first sentences of "Never look back," Ken C. Pohlmann's "Digital Horizons" column in the January 2005 *Sound & Vision*. The rest of the column seems to be saying that HD-DVD is *really* being introduced to make us all re-buy all the DVD movies we purchased (and because it's even more ludicrous to file-share than DVD), not because it offers four times the picture quality. But really, after "truisms" as absurd as those stated here, all I could think of is another truism: "It's tough to write satire these days."

➤ The San Francisco *Chronicle* for January 23, 2005 reprinted a charming little piece by Elizabeth Large, originally in the Baltimore *Sun*: "Even the rat race needs a slow lane." It's about multitasking and effectiveness, citing a poll and study in the new *Scientific American Mind* and relating it to Carl Honore's *In praise of slowness*." The poll found that 90% of Americans multitask—and that 60% of us "said they felt as though they were getting done." The article "concluded that multitasking doesn't work very well, unless you're doing something routine like walking and chewing gum." Dr. Barry Gordon says we need to slow down in order to do any real thinking—and, thus, Honore's book. I don't know about the various "slow movements," but I'm becoming increasingly skeptical about multitasking as a way to improve overall effectiveness. I suspect that, too often, it's a way to do several things badly—something I've noted before. In my own life, I find that I no longer play music while I'm writing; because I almost always care (at least a little) about the music I listen to, it distracts from the writing. Your multitasking may vary, of course.

- "Dot.life" from the *BBC News Magazine* online had a brief piece by Paul Rubens on November 29, 2004: "Sound of music." It leads off with a discouraging question: "When did you last read anything about an MP3 music player that mentioned the quality of the sound it reproduces?" Rubens goes on to say that advertisements and reviews talk about size, storage capacity, types of music files, even colors—"but chances are that sound quality won't get a look in." (British, remember, and the article said "colour.") He suggests that the answer is "quantity turns out to be more important than quality" (but see LIBRARY STUFF in this issue, specifically the notes about DPE). He believes most people simply don't care about quality any more and that "what music lovers want today more than anything else is music—and lots of it." He applies the same dystopian assertion to video, where low-quality DivX is a popular alternative to DVD. Finally, he suggests that the slow comeback of "vinyl and valves" (tube equipment) may mean a renaissance of serious music listening. Well…I think the "vinyl and valves" stuff deals with something other than high fidelity sound reproduction (but I could be wrong). I *know* that a lot of people care about quality sound reproduction—and that most people have never heard true high fidelity reproduction, MP3 or otherwise. Fortunately, I had a ready answer to the leading question: The most recent group review of MP3 players in *PC Magazine* considered reproduction quality as part of the evaluation. Some sources *do* care.
- While the "evolution stickers" in Cobb County School District were overruled this time around, similar inane attempts will certainly arise again. There's a great page of alternative stickers—including the original, with exactly that wording—at www.swarthmore.edu/NatSci/cpurrin1/textbooksdisclaimers/ The page includes 15 stickers in all; it's also available as PDF to print out on actual sticker stock (or you can blow up a single sticker to make custom t-shirts—or buy one from CafePress). Remember the original? "This textbook contains material on **evolution**. Evolution is a theory, not a fact, regarding the origin of living things. This material should be approached with an open mind, studied carefully, and critically considered." Among the others: "This textbook contains material on **gravity**. Gravity is a theory, not a fact, regarding a force that cannot be directly seen…" (and the last sentence of the original). (Similarly for heliocentrism, the theory that the Earth orbits around the Sun; plate tectonics; and special relativity.) Better: "This textbook claims that evolution is not fully accepted by scientists because it is just a theory. The author hopes to confuse you into equating 'scientific theory' with 'cockamamie theory.' To read a short blurb on what a scientific theory is, go to http://wilstar.com/theories.htm" Great stuff.
- A little report in the January 2005 *Computer Shopper* says that MP3 is still the "overwhelming favorite music format of file traders," but that it's slipping. Given the popularity of the iPod, that's scarcely surprising. What's unfortunate here is sloppy or ignorant reporting: "The percentage of songs in MP3 format in people's digital-music collections has slid down to about 72 percent, from about 82 percent a year ago." Since the piece goes on to say that Windows Media files make up 20% and Apple's AAC format 4.3%, there's a major oversight here: What about all that digital music in the form of CDs? Or does being on a physical carrier make it no longer digital?
- Since I'm picking on *Computer Shopper*, here's another oddity. Each month, they have a "buying advisor" feature where a reader needs a particular functionality at a particular price and they consider the best alternatives, finally offering a recommendation. This time it was an ex-teacher who aspires to be a writer and wants an "easily totable device" to write in her spare moments; she wants to spend no more than $1,000. The discussion is interesting; cheap notebooks tend to be heavy, while light notebooks tend to be expensive—and some compromise units lack optical drives at that price. So far, so good—but the final recommendation is NEC's MobilePro 900c, an oddball 1.8lb. "half-VGA" semi-PDA with a "typeable" keyboard (but is it truly touch-

typeable?). It uses Windows CE and MS WordPad. And, like any other PDA, its CD-ROM drive is…hmm, I don't see an optical drive anywhere in the specs. Any more than I would for any other PDA or pseudo-PDA. I suppose consistency is a bit much to ask for.

## Perspective

# Wikipedia and Worth [Revisited]

*Cites & Insights* 4:12 (October 2004) included PER-SPECTIVE: WIKIPEDIA AND WORTH, based on a test in which Alex Halavais changed a baker's dozen worth of Wikipedia entries and described the results—and on a whole bunch of commentaries for and against Wikipedia. Site statistics show that 4:12 was downloaded by more unique visitors than any other issue during 2004, so most of you may already have read that essay.

The controversy over Wikipedia didn't stop. There's been enough action since then, in the web community and on library lists, that it makes sense to revisit the scene. I am quoting from library lists in this case, albeit selectively—noting that these lists are open for anyone to join and review the archives.

## Opening Salvos

On November 15, 2004, Robert McHenry (former editor in chief of the *Encyclopædia Britannica*) posted "The faith-based encyclopedia" at *Tech central station* (www2.techcentralstation.com). He discusses an early internet discussion proposing the Interpedia, "a reference source for people who have connectivity to the internet." McHenry says the discussion group "generated a great quantity of writing, none of it encyclopedic in nature"—discussions of necessary software, how to attract contributors and organize teams, how to allow editing but prevent "unauthorized alteration," and "rhapsodic explanations of why the Interpedia, as a noncommercial and collaborative product, was ipso facto superior to all existing encyclopedias, all of which were published for *[shudder]* profit and all of which had their origin in *[shudder]* print."

He concludes that, although some participants said "how do we start," the discussion eventually petered out, "in part because some real encyclopedias developed Internet presences, and in part because the volunteer nonleaders of the ungoverned, unstructured project truly did not know where or how to begin."

A decade later, however, "the Wikipedia project is flourishing." As of November 2004, some 30,000 contributors had written 1.1 million articles in 109 languages, including more than 382,000 "pages that were thought 'probably' to be encyclopedic articles" in the English version. (The Manx Gaelic Wikipedia only had three articles—but the Klingon version had 48.)

Impressive effort, attracting positive press. As he notes, the founders built (or found) the software (wiki), attracted contributors, generated buzz—and created the needed background hierarchy to keep things going. "The question is, however, just what have they created?" He cites the FAQ to see what they *intended* to create:

> Wikipedia's goal is to create a free encyclopedia—indeed, the largest encyclopedia in history, both in terms of breadth and depth and also to become a reliable resource.

McHenry believes that the order of adjectives is significant (reliability is less important than free, broadest, and deepest); I'm not inclined to parse the sentence that carefully. He goes on to note that the statement is "like everything else on the Wikipedia site…editable, by anyone."

McHenry cites other stuff about collaborative editing, noting that the process is claimed to allow Wikipedia "to approach the truth asymptotically." Then McHenry starts to trash the claims and the methodology. Ed Felten did this earlier (but much less emphatically), suggesting that Wikipedia's editorial process will result in a "random walk" around some plateau of quality that may or may not be as good as (or better than) a traditional encyclopedia, but is unlikely to be the "highest degree of accuracy" that Wikipedia's advocates assert as a reachable goal.

McHenry doesn't buy it. He mixes the "moist and modish notion of 'community,'" "vague notions about information 'wanting' to be free," and "journaling" as part of education. He knows it's nearly impossible to carry out a thorough assessment of an encyclopedia's quality—so he looks at one article in Wikipedia as a case study. He's not impressed, asserting numerous typographic, styling, grammatical and diction errors and calling it a "C [high school] paper at best"—despite more than 150 edits. He also says earlier versions were better written—"the article has, in fact, been edited into mediocrity." He concludes:

The user who visits Wikipedia to learn about some subject, to confirm some matter of fact, is rather in the position of a visitor to a public restroom. It may be obviously dirty, so that he knows to exercise great care, or it may seem fairly clean, so that he may be lulled into a false sense of security. What he certainly does not know is who has used the facilities before him.

Before commenting further, I should point out that *Tech central station* is a distinctly right-wing website. Its slogan is "where free markets meet technology." It's also fair to assume that a former encyclopedia editor will be inclined to favor print encyclopedias. Is the *Britannica* free of error? Certainly not—and in reading a given article, I may or may not have much idea who's "used the facilities before me" (whose editorial hands have touched and changed the writer's contribution, if I know who the writer is or why I should trust them). It would be easy to ignore this essay as sour grapes.

Four days later, Jason Scott posted "The great failure of Wikipedia" to one of his weblogs, *ASCII by Jason Scott* (ascii.textfiles.com). His essay runs seven double-spaced pages, followed by four pages of comment-and-response as of January 4, 2005.

Scott researches computer history: That's the focus of this weblog. He begins:

> I have now tried extended interaction with Wikipedia. I consider it a failure. In doing so, I will describe why, instead of just slinking off into the night on my projects. Maybe it will do some good. Maybe it will not. I'm sure, at the end of the day, there must be hundreds like me at this point. Burned, slapped, ejected from the mothership for not following the rules, no matter how intricate and foolish. Let me at least go with some smoke.

> The concept of Wikipedia is a very engaging and exciting one, especially to someone like myself who spends an awful lot of time collecting information and then presenting it to people. Normally, the work I do is the work that's done. That is, if I don't give much attention to a specific section of my sites, that section will stay static, even if it's in need of improvement. This is not very enjoyable. In collaboration, you will put your tools down for the night, and when you wake up the next morning, more work is done. This is very exciting, very enjoyable. It's why people work in teams in the first place.

He describes his own (usually solitary) working habits, the enticement of Wikipedia, and running into people who regard it with "a near-fanatical aspect." He

wanted to get more people involved in researching computer history, e.g., the history of AOL. So he signed on to Wikipedia and started some work. At this point in the discussion he tells Wikipedia workers that he's going to make them angry: "What I am doing is trying to stop people from working on Wikipedia with the idea that they're accomplishing good."

Scott *does* do history, so he brings in the Usenet FAQ (one of the earliest FAQs), which succeeded in part (he says) because of "a lot of collaborators but a short list of people maintaining it." He notes *The Mythical Man Month*, a classic computer-related book, which concludes that adding people to a project that's behind will usually cause it to fall further behind. Going off to one side, Scott uses IMDB user comments to suggest, "A low barrier to entry leads to crap." All of this may be relevant to Wikipedia.

His primary criticism: Wikipedia has "a small set of content generators, a massive amount of wonks and twiddlers, and then a heaping amount of procedural whackjobs"—and the larger groups mean content generators "have to become content defenders," defending their expert text against changes from a larger group of people with no expertise in the subject. He also hates the "Neutral Point of View" espoused by Wikipedia: "Like wikipedia itself, it is a great idea in theory. In application, of course, it turns into yet another hammer for wonks and whackjobs to beat each other and innocent bystanders."

> I'm sorry, but content creators are relatively rare in this world. Content commentators less so. Content critics are a dime a hundred, and content vandals lurk in every doorway. Wikipedia lets the vandals run loose on the creators, while the commentators fill the void with chatter. It is a failure.

Scott goes on to say that he might make most of his works available "essentially for free to the public" but will not allow Wikipedia to use them. He refutes the comparison between Linux and Wikipedia is false and notes that there are other wikis out there with higher barriers to editing to "ensure that the person who is going to undo your hours of work with a few mouse clicks is at least, from some relatively objective standpoint, vaguely entitled to do so."

There was feedback, although I'm guessing that most flames showed up elsewhere. What shows up here is a civilized back-and-forth among Scott and four others. Notably, Scott *agrees* with "peterb," who says Wikipedia is a success *for users* even if not for

some creators: "You are correct; from the outside, to someone who is looking for basic information, a lot of Wikipedia will be 'good enough.'"

Five weeks later, on New Year's Eve, Larry Sanger published "Why Wikipedia must jettison its anti-elitism" on Kuro5hin (www.kuro5hin.org)—and all hell broke loose, or at least the kind of hell that happens on Kuro5hin and slashdot. Sanger's five-page "op-ed" (labeled as such) generated 397 comments as of January 27 (312 of them by January 5), adding up to more than 62,000 words (considerably longer than my latest book). Some of these people have *way* too much time on their hands.

Sanger co-founded Wikipedia but has since left the project. He says up front, "I know Wikipedia is very cool. A lot of people do not think so, but of course they are wrong."

So what are Sanger's issues? "First problem: lack of public perception of credibility, particularly in areas of detail." He's not saying Wikipedia is unreliable—but that it's *perceived* as inadequately reliable "by many librarians, teachers, and academics." Saying "but it gets used a lot" does nothing to negate that problem: "people use many sources that they themselves believe to be unreliable, via Google searches, for example." He goes on to point out the benefits of credibility—and to suggest there's a *real* problem with credibility in specialized topics outside the interests of most Wikipedia contributors. (He suggests comparing Wikipedia's philosophy section to the *Stanford Encyclopedia of Philosophy*, as one example.)

"Second problem: the dominance of difficult people, trolls, and their enablers." See Scott's article, above. "Far too much credence and respect accorded to people who in other Internet contexts would be labeled 'trolls.'" He thinks this is a generic problem with unmoderated Usenet groups that's infected Wikipedia—although he notes that Wikipedia takes steps to control the problem in its most extreme cases.

"The root problem: anti-elitism, or lack of respect for expertise." "[A]s a community, Wikipedia lacks the habit or tradition of respect for expertise." He comments on his efforts to overcome anti-elitism (including snubs and disrespect of expertise) and the consequences of the current situation. Experts with relatively limited time and patience don't participate, because they have to defend their positions and are shouted down if they complain about whack jobs.

Sanger believes that Wikipedia's openness does not *require* disrespect toward expertise. He believes Wikipedia would be much better if more experts contributed. He anticipates a "more academic fork of the project" at some point.

## Responses and List Comments

People piled on the op-ed right away. That's typical of high-energy threaded sites. In the process, they revealed the range of attitudes that make Wikipedia possible—but that may hold it back from being what it could be. Originally, I only looked at those comments that show up at default settings of *Kuro5hin*, less than 25% of all comments by number but still more than 25,000 words. (When I went back to download the whole set of comments, I found myself unwilling to read through the full 400.)

Wikipedia supporters and fanatics—not always the same people—exhibited a whole range of defensive mechanisms, from variants of "it isn't *meant* to be the equivalent of a print encyclopedia" to "if something's wrong, it's up to you to fix it," with strong doses of anti-elitism, anti-intellectualism, and "You left the project, therefore don't get to say anything about it" mixed in. Several people agreed with some or all of Sanger's commentary. Several people noted that their own work had been used in Wikipedia without attribution—and, in one case at least, their attempt to provide such attribution was deleted.

Here's just a few of the comments:

➤ "I once considered submitting some articles to Wikipedia on the topics I am expert on…but after I looked through the site I realized I didn't have the time or patience to deal with the fools that inhabit the project." (dharma)

➤ "…any problem wikipedia has can probably be traced to some sort of residual elitism somewhere: amongst a contributor, or what a questionable rule promotes in contributor behavior" (circletimesquare)

➤ "I've read Larry Sanger's Wikipedia articles: they're mostly crap… The simple fact of the matter is that credentials are no guarantee of competence. A huge percentage of academia is filled with professional bullshitters, and they seem to be pissed off that Wikipedia is doing an end-run around their nonsense." (Delirium)

- "Hey, Larry—why are you writing about Wikipedia when you yourself proved incapable of sticking with the project?... Either go back and contribute, or stop talking smack about a project you're no longer involved with." (grendelkhan)
- "I love Wikipedia, but I'd also like to have a reviewed-by-academics section. Why? So I can finally cite it in my damn essays without getting "is this source reliable?" written there, that's why." (reklaw)
- "Larry's comments betray a complete ignorance of the project… I am not anti-exertise [sic] in any fashion…" (from its current head)
- "Wikipedia seems, to me, to be society's present best chance to rise above the current acedemic [sic] principle of respecting only the work of someone with a PhD, and encouraging high standards of research and writing among society at large." (J T MacLeod)
- "Wikipedia is just fine, and the more the 'experts' squawk and complain, the greater the evidence that it is so. This is the *Age of Participation*, and self-correction will ultimately win out, because experience, not expertise, is the new authority." (Also lots of stuff about postmodernism or PoMo and why expertise no longer matters.) (xnuzboss)
- "Certainly, if what Wikipedia is after is 'official' recognition of it being an accurate and reliable source something needs to be done. But, I don't think that is what is sought after at all… The content of articles do not necessarily have to have a one-to-one correspondence to any particular version of reality." (cdguru)
- "No encyclopedia is reliable, Wikipedia included. And no encyclopedia should be used for anything but a starting point…" (vadim)
- "We must explore what networked knowledge can give not mimic what an eatablished [sic] encyclopedia already can do. I don't care if conservatively thinking minds can't accept wikipedia." (drquick)
- "I question one of your most basic premises: that Wikipedia works… I will assert…that Wikipedia simply fails to achieve the status of 'Encyclopedia.' It is complete, comprehensive

- and indexed, but it is not accurate or readable." (cjames53)
- "Unfortunately, in today's complex world there is not simply a single objectively true point of view…" (MoebiusStreet)
- "One should never fully trust any source regardless of who (expert or not) 'approved' it or did not approve it. This is the real beauty of wikipedia, it teaches us to always question, to seek out other sources, compare and contras [sic] and be critical." (thehero)
- "Summary of 1,000 'arguments': Third party observer: 'One can produce a reference work of better quality than Wikipedia currently does by doing X.' Wikipedia cult member: 'Doing X will not produce a perfect reference work, therefore, it's of no value. And anyway, if there's an error in Wikipedia, it's your own f[**]ing fault for not fixing it." (Estanislao Martinez, with three-letter modification in quotation)

Just glancing at the first few dozen of several hundred additional comments shows a pattern of even less polite and thoughtful comments than these.

*Weblogs and Wired News*

Karen Schneider wrote "Wikipedia's reality check," on January 1, 2005 at *Free range librarian*. She discusses Stanger's post and some responses as well as Scott's article. She's intrigued by the fact that Sanger helped develop Wikipedia, calling it "the triumph of hope over experience."

Schneider notes that her operation, Librarians' Index to the Internet, *does* use Google—but only to find potential resources. LII doesn't point to Wikipedia articles because they're too fungible: "we don't have the time or resources to constantly monitor the Wikipedia entries or dig through long, heavily trolled discussions to verify the authority of a resource."

In the end, Schneider pretty much dismisses Wikipedia:

> Good encyclopedias already exist. Wikipedia is fixing a problem that isn't there, and in doing so, with its endemic, unsolvable, inherent problems, it is revealing the naïveté of its creators and the predictable characteristics of unmanaged electronic territory.

If you take Wikipedia as a direct equivalent or competitor to a print or CD-ROM encyclopedia, I'd agree—and it's pretty clear that that's how Wikipedia's creators view it. But I think there's more to it:

Wikipedia *does* serve a different and (potentially) valuable purpose.

j Baumgart posted "Wikipedia and a controlled vocabulary at *j's scratchpad* on January 5, 2005. She's noticed problems with many of the linking terms: They're inconsistent, which leads to problems. She's asked whether "Wikipedians" patrol the encyclopedia looking for name and link issues. She suggests that a controlled vocabulary might help. Baumgart is, after all, a librarian.

Clay Shirky lost it (in my opinion) in a rant at *Many2many* that included this comment:

> Of course librarians, teachers and academics don't like the Wikipedia. It works without privilege, which is inimical to the way those professions operate…. You can see the reactionary core of the academy playing out in the horror around Google digitizing books held at Harvard and the Library of Congress… The physical book, the hushed tones, the monastic dedication, and (unspoken) the barriers to use, are all essential characteristics of the academy today." (Quotations from *Godwin's law*, a January 5 posting. I'm not quoting directly from *Many2many* for the usual reason: as with a number of other weblogs, it is nearly impossible to print more than the first page of any posting, and I've pretty much given up.)

I was astonished by Shirky's entry when I read it. His flat-out attack on (all?) librarians, teachers and academics seemed extreme. In later postings, he's dismissed professional taxonomies as being of *any* use on the basis that 'folksonomies' are cheaper for *some* uses, so maybe this wasn't an isolated incident.

Danah Boyd also did a *Many2many* commentary, "On a vetted Wikipedia, reflexivity and investment in quality." She doesn't believe Wikipedia should be ignored but also doesn't believe it will ever be a (traditional) encyclopedia. She uses the page on "anomie" as an example of the problems with Wikipedia entries. She agrees with Shirky that it's a "system not a product" and "value[s] it intensely" but doesn't buy into the religion that gives it authority "simply because it is open-source." She also notes *why* academics deserve some credit for their expertise.

A January 10 *Wired News* piece by Daniel Terdiman, "Wikipedia faces growing pains," brings another face of Shirky: A question as to whether Wikipedia's methodologies scale well. He quotes Danah Boyd saying that she finds Wikipedia "an exceptionally valuable tool," but notes that non-technology entries may not be very good. The article quotes Larry Sanger and also Jimmy Wales' assurance that Wikipedia is healthy: "It is increasingly being cited and relied upon in news by academics, librarians and researchers." (Wales is president of the Wikimedia Foundation, home of Wikipedia.) That's astonishing if true…but Wales goes on to dismiss the notion that Wikipedia "or any encyclopedia" should be a top-tier reference source. He also notes that Wikipedia will be frozen at some point, with a new one starting then.

## Web4Lib

At least 24 postings related to Wikipedia turned up between January 3 and January 9, 2005, starting with "Co-founder of Wikipedia talks about problems" and changing to either "Generation shifts and technology" or "Seeing ourselves as others see us."

Alain Vaillancourt, who's a participant, said no academic group would take on improving Wikipedia: "Wikipedia is far from being an encyclopedia in any sense of the term. It is a better organized outgrowth of communal blogs and nothing more. The best solution would be to drop 'pedia' from the name and put in something else." Brad Eden (who knows Larry Sanger) agreed: "There will never be an academic interest in this work, because it just doesn't have any type of adequate peer review or group that can adequately review the authority and accuracy of what is presented or 'published.'"

Lars Aronsson raised a bunch of controversial points: "First we must agree that Wikipedia is needed, or this discussion becomes pointless. Second, improvements over what Wikipedia is today must be possible to implement… That the web needs a free encyclopedia…is shown by Wikipedia's outstanding popularity. Many people are buying the concept, even if some librarians aren't… If it were as bad as you suggest, it would drop off the [Alexa list of most popular websites] pretty soon, wouldn't it?... Innovations always appeal to new categories of users. Wikipedia is primarily for Internet users, not for traditional users of encyclopedias."

Andrew Mutch questioned the final statement: "What's the distinction between the two groups and on what basis do you make that claim[?] It implies that people who use encyclopedias in the library won't get online to use a similar resource. If so, I would be interested in seeing anything that shows that to be true." Bill Drew took issue with Aronsson's first assertion and noted, "Just because people use it does not mean it is needed. It could mean that they

are unaware of better resources…. Until it becomes a resource created by experts and verified in some way, it will always be suspect in my view." He also took issue with the "generation shift" claim. Drew, like me, has been using the internet and computers since very early days—and doesn't consider himself part of the "non-computer" generation.

Ross Singer agreed with a Drew point I didn't quote (people using whatever's easiest) and said it's why libraries and librarians "are struggling to get patrons to use the clumsy resources that they pay so much to have access to… What makes [Wikipedia] special (and therefore, more valuable than academic and 'vetted' sites, IMHO) is that people are actually *using* it." [I would note that people use "vetted" paid library online resources by the *millions* every month and print resources by the tens of millions every week: Maybe not as much as Ross or I would like, but usage statistics for subscription resources belie any claim that the resources aren't being used.]

Vaillancourt suggested "traditional paper encyclopedias" are inaccessible to most people because there aren't libraries in every neighborhood. "So even a badly flawed effort like Wikipedia fills a certain information need." [Nobody pointed out that you can buy a CD-ROM encyclopedia for $30 or so—and anyone with the equipment to use Wikipedia almost certainly has the equipment to use such an encyclopedia, unless they're using a library terminal nearby one of those paper encyclopedias.] He also asserted that there are two Wikipedias: general articles from the general web community, Open Source and related computer articles—with lots of depth and considerable expertise—from the Open Source community.

Aronsson returned, pushing the "generation shift" concept (I call it KTD, and I'm with Mutch in questioning its general applicability). He asserts that Karen Schneider's preference for a traditional encyclopedia and statement that Wikipedia isn't needed "comes from her mindset where she looks for an encyclopedia, rather than linkable information on the web." This is the closest I've seen anyone come to calling Karen a technophobe or traditionalist! Aronsson says that, if he's writing on the web, he'd just link to Wikipedia for more information "because I am writing on the web, my readers are on the web, and Wikipedia is on the web." He also says, "fans of Wikipedia will tell you that it is already an encyclopedia, and fans of encyclopedias are likely to disagree."

Karen S. returned at this point, flatly disagreeing with Aronsson's "First we must agree" assertion—"There's no rhetorical support for that position in this discussion"—and responds to the "needed because popular" claim by asking, "Do people need cigarettes and SUVs?...There's no strong relationship between what people need and what they use… There is a need for high-quality, freely-available information, and I'm all too aware of the gaps out there. I wish some of the energy behind Wikipedia had gone into advocating for the retention and restoration of important information resources that have been removed from the Internet."

Aronsson backed down just a little: "Perhaps 'need' is a word I should avoid…" His response to Bill Drew's argument that use of Wikipedia is not a classic generation shift is to beg the question: "Then, could you please describe how classic generation shifts work?" [Sorry, Lars, but if I say that *Lady Chatterley's Lover* is *not* pornography, the response "Then, could you please describe what pornography is?" isn't particularly useful.]

Fiona Bradley admitted that she uses Wikipedia and other wikis. "For me, as a librarian working with a very small reference collection and incredibly short deadlines, Wikipedia provides a launching point [for further research]… Most publications are not peer reviewed, including books, journals and conference papers." There's more to this sensible commentary.

Ryan Eby, who "would probably be considered part of the 'generation shift' away from print media," chimed in on why Eby uses Wikipedia. "For starters, almost no teachers…will allow encyclopedias to count as a source… I'm not the kind of person [who] trusts any one source, be it an encyclopedia, website or book by 'experts.'… [Main reasons Eby uses Wikipedia:] Constantly updated… References: Most of the information…have…a good amount of references…I use Wikipedia as a starting point… Open access… Breadth of coverage… Multiple viewpoints… In general I wouldn't call Wikipedia a[n] endpoint in research but as a starting point I probably couldn't ask for much more."

At about this point, the prevailing subject head changed to "Generation shifts and technology" and some posts didn't relate directly to Wikipedia. Bernie Sloan wasn't ready to buy the "generation shift" concept entirely. "The idea seems to revolve around each new human generation being more technologically

adept, or integrating technology into their lives more completely, than the preceding generations. But I see evidence that this isn't so cut and dried." One of his twenty-something kids is really wired (or "unwired"); the other went six months without internet access "and it didn't seem to bother him a bit." One of his brothers checks email once every nine months—and his 80-year-old mother has "been an avid home computer user since the days of the Apple IIe."

> I guess my point is that every generation has members who take to technology like a duck to water, and every generation has members who aren't particularly technologically adept, and who couldn't care less that they aren't.

Fiona Bradley (who's 26) agreed: "I have seen assumptions made about younger generations…over and over again with little substance to back them up." She notes that these assumptions mean that kids aren't getting training in technology. It's assumed they'll know it intuitively.

Steve Cramer noted, "It's easy, imo, to get into stereotypes about technology use and age." When he trained for the local freenet, "half the folks" at users meetings "were retirees, or at least grey-haired; there was also a large number of young adults and a few high school kids, but not too many folks in between…. Ten years later, I bet those older netizens give tips to their grandkids…on searching the web or customizing their browsers." (Art Rhyno later noted that at one point, in Canada, senior citizens represented the fastest-growing segment of internet users.)

Aaron Dobbs moved the "generational shift" idea in a different direction—one I mostly agree with, despite my constant poking at KTD oversimplification. He notes that we keep getting *more* technology and that descendants of early adopters are likely to seem more adept and, themselves, also be early adopters—but that doesn't mean all kids are technophiles.

"ChuckO" flatly denied all evidence that every generation has some members who aren't particularly technologically adept. In a grand universalist statement, he asserted: "This may be true of the older generations, but the younger generations breathe technology…" He bases this *absolute universal* claim—he explicitly *denies* that there are *any* young people who don't "breathe technology"—based on working with young people and looking at the demographics of Friendster. That, and the fact that his 70-year-old mother refuses to touch a computer. Simple, isn't it:

ChuckO has *proved* that all young people "breathe technology" and that all old people avoid it entirely.

Ryan Eby wasn't buying this at all. "I still think it's true for most generations (not being inherently tech savvy). While [tech savviness] may be more prevalent with younger generations, it's far from 100%... While more and more are growing up with technology, it still can't be 'presumed.'" He notes that the messages have *demonstrated* that plenty of people, at all ages, "defy the norms."

Postmodernism had to arrive at some point—and, arriving after Shirky's jumping-the-shark moment and broadside against librarians, there was also a new subject, "Seeing ourselves as others see us." Larry Campbell's take: "What I find most interesting about the wikipedia as a web phenomenon is the way in which it affects the whole idea/ideal of 'authority' and 'expertise.'… Perhaps the wikipedia isn't so much imitating the authority of the encyclopedia but rather inducing a more critical awareness of the nature and limits of 'authority' as such."

The threads stopped around January 9; that usually happens with a hot topic on Web4Lib (and most other lists). The last post I saved was from Lars Aronsson, who clarified that he "wasn't talking about human generations at all, but of generations of technology replacing each other"—pointing us again to Clayton M. Christensen's *The Innovator's Dilemma*. Aronsson's examples of media shifts within computing didn't make much sense to me, but I haven't read the book—and I don't buy the idea that the web represents a *technological* generation shift over print resources. It's rarely that simple. It's certainly not that simple when comparing Wikipedia to, say, *Encarta*. Aronsson also felt the need to slap librarians around a bit: He says Wikipedia's fans "are not reference librarians who earn their living by navigating the paper-bound information space." Think of all those reference librarians out there who ignore online resources… I don't know of any. Do you?

### Publib—and coming full circle

While I may have ignored some, I did print off three Wikipedia-related posts on Publib, all on January 11, all responding to Chris Ely, who was curious to see if anyone was using Wikipedia for reference, "even if it's just a starting point to find other materials. My concern is that 'anyone can edit' and I'm curious about how accurate it can be without someone overseeing the edits."

Elizabeth Thomsen is a fan: "[T]he beauty of the Wikipedia…is that instead of *someone* overseeing the edits, *everyone* is overseeing the edits. I was skeptical when I first heard about this… I was surprised at the quality of most of the articles I read… Exploring more, I learned why the Wikipedia works as well as it does [She goes on to note the primary mechanisms.] It's really quite an amazing community, not at all the anarchy you might be imagining from the scary words—'anyone can edit'!"

Jim Deane started with the standard and appropriate caution: "I think you must be careful with any information source." He uses Wikipedia and has contributed to the effort. "My area of specialty is physics, and the articles I have read on physics topics have generally been first-rate… I would look for verification sources (books, journals) for most information in Wikipedia… If you look at the Wikipedia content controls, you can see that there are actually some ways that abuse can be curtained…"

Chris Rippel is impressed with the currency of Wikipedia, noting its post-election coverage and, more recently, its article on the 2004 Indian Ocean earthquake. Rippel preceded that by recommending that Chris Ely read "Walt Crawford's Perspective on Wikipedia [in *C&I* 4:12]."

## Conclusions

Which brings us back to that essay. Here's what I concluded last October:

> Wikipedia is certainly not worthless. Wikipedia is also not automatically better than a traditional encyclopedia because of the community of writers. I would tend to use Wikipedia entries as starting points, to be used on a "Trust but verify" basis. But isn't "trust but verify" the base heuristic for almost all resources, traditional or new?

> My assumption is that lots of specialists have contributed good work to Wikipedia, particularly in areas related to the web and digital resources. My assumption is also that some Wikipedia content is faulty, biased or wildly incomplete. In the latter case, I'd make the same assumption about a traditional encyclopedia, up to and including Britannica. Personally, I doubt Wikipedia will "eclipse" traditional encyclopedias (note that Britannica is once more available in print form), just as I doubt that weblogs will replace newspapers or that econtent will sweep away print media. Another comparison may be apt: While Encarta may have doomed Funk & Wagnall's (and incorporated it at one point), it hasn't doomed

Britannica—or vice-versa. Different forms, different media serve different people with different needs.

> I did my own tests—only of Wikipedia, and if I was going to compare it with a traditional encyclopedia I'd probably use Encarta or its ilk. My results were mixed. Some entries were very good and reflected considerable expertise (I was particularly impressed by the network of entries on lossy compression technologies), while some were dry as dust and gave no flavor of the subject they dealt with. I didn't catch obvious errors, but I wasn't really looking.

> What is an authoritative resource? How about articles in refereed STM journals? Are they automatically trustworthy because they've been through that rigorous peer review process? Not really, as candid observers within the field will tell you. Very few observers argue with the cynic's view of peer review as it applies to authors with an axe to grind: Peer review doesn't determine whether an article will be published, only where it will be published.

> There was a dustup on the SPARC Open Access Forum about quality and inclusion in indexes. Implicit in the dustup—which I won't review since it's pretty specialized—is the notion that a great many peer-reviewed journals are full of crap. Sometimes that's implicit (to a knowledgeable observer) from the name of the journal. Sometimes you need to be aware of the standards of the field to know which journals are, in the words of one observer, essentially vanity presses.

> This is a tangent, but I sometimes wonder about the dissociation of article and journal that appears in online aggregation and OAI. Specialists will have their own mental lists of first rank, second rank but trustworthy, and FoC [see previous paragraph] journals; they'll use those mental filters to judge the worth of new articles. But what about non-specialists? Are those filters readily available? Who will tell a person looking for health guidance which peer-reviewed journals to avoid at all costs?

> I think there is a connection to Wikipedia. Alternative publishing does not imply lack of worth. Traditional publishing isn't an automatic indication of worth or veracity. If the key is "trust but verify," we need better ways to verify likely worth and probity. I assume today's librarians are finding such ways, and finding ways to communicate those methods to the rest of us. At least I hope so.

I don't see much reason to change those conclusions.

Why did I spend six thousand words on this brouhaha? Because the process is itself interesting. Because I see a narrative arc here, and I love story tell-

ing. Because even the top-level Kuro5hin comments are noteworthy—if nothing else, in showing how extreme advocates for a cause or position can claim to undermine their opponents without actually responding. You say Wikipedia doesn't stand up to scholarly scrutiny and you can prove it? Well, then, it's not really intended to *be* an encyclopedia, in spite of the site's assertions. You say the facts are wrong sometimes, or that they don't agree with established authority? It's a new era, and authorities don't matter: Wikipedia represents *some* truth, even if not your tired old reality-based truth. You say Wikipedia needs improving? Well, unless you're spending all your time making improvements, you're not allowed to comment. To most of which I would respond in words that won't appear here, since I try to keep the swearing to a minimum.

But I also get a little edgy when Wikipedia is attacked for not being an exhaustive scholarly resource. *No* encyclopedia is (or should be) an exhaustive scholarly resource. Encyclopedias should *always* be starting points.

I was more favorably impressed by the threads at Web4Lib and Publib than by the interminable comments at Kuro5hin. Maybe I shouldn't be. LISNews demonstrates that you can use slashdot's software (or a variation thereof) without descending into slashdot's endless flamewars and ignorant hyperbole (perhaps overstating the case, but that's my experience). Web4Lib continues to astonish with its usually polite, usually thoughtful, *unmoderated* postings. Even ChuckO's extreme position was stated calmly, and nobody called Aronsson a fathead or worse. (Nor should they have. I may not agree with what he Aronsson saying, but the comments were thoughtful and carefully worded.)

Universalist arguments are usually nonsensical, including KTD/generation shift arguments. "We all" do very little based on age or similar broad-brush characteristics (including being human), except breathe, eat, and eventually die.

Arguing for Wikipedia's superiority by shifting the definitions suggests weakness in the case for Wikipedia. I believe Wikipedia's supporters err in pushing the "better than print" argument. Better they should make Wikipedia the best online resource it can be (which may mean showing a little more respect for authority—I really don't know!), while granting that traditional encyclopedias also have considerable

worth. Maybe "pedia" as part of the name does more harm than good—again, I'm not sure.

Maybe it's human nature (for some humans, not all) to advocate your own preferred solution by putting down alternatives rather than by showing the virtues of your choice. That's sad if true. Wikipedia can do just fine. So can *Encarta.* So can *Britannica*, back in print and still in digital form. And so, to be sure, can all of those books, journal articles, "vetted" websites and primary sources that encyclopedias of any nature should lead us to.

# The Good Stuff

"25 years of technology," *Computer Shopper* 25:1 (January 2005): 91-5.

It's really only two pages (page 91 is an introductory paragraph, 92-93 a Dell ad), but it's still an interesting brief timeline of the quarter-century since *Computer Shopper* started. (Which, itself, was five years after the first modern PC, the Altair, was introduced.) Some tidbits: The first 1GB hard disk came out in 1980, from IBM, "as big as a refrigerator." The first million-selling computer came out in 1981—Commodore's VIC-20. (That's also the year IBM introduced the IBM PC.) ARPAnet switched to TCP/IP in 1983, which thus counts as "the start of the Internet." The Mac came out in 1984—and Michael Dell started selling computers. PS/2—remember PS/2?—came out in 1987; Windows 3.0 (the first usable Windows) in 1990; Linux in 1991. The first White House website emerged in 1993, as did the Pentium and Mosaic. CD-RW didn't happen until 1996—and DVD went on sale in the U.S. in 1997. And in truly paradigm-shattering news, the Segway Human Transporter appeared in 2001: Cities have never been the same! Or, wait…

Anderson, Chris, "The zen of Jeff Bezos," *Wired* 13:1 (downloaded January 7, 2005).

What makes this brief interview noteworthy is where Amazon's founder stands these days on the inevitable triumph of Amazon over physical bookstores and the like. Maybe it's not new; maybe digital triumphalists overinterpreted Bezos' plans. Anyway, here's Bezos' estimation of how retail sales will settle out: "I think online ultimately will be 10 to 15 percent of retail. The vast majority of retailing will stay in the

physical world…" He doesn't exempt bookstores, noting that physical bookstores are one of the "third places" people gather. "We humans are a gregarious species; we like to mingle with other humans."

Bezos notes that Amazon sales are disproportionately weighted toward harder-to-find titles compared to the book industry as a whole and claims that the collaborative recommendation system creates demand for hard-to-find products. He notes that Search Inside the Book has apparently increased sales of those books—even cookbooks and reference titles, where one might think the snippet would be enough.

Curiously (or not), Bezos doesn't expect print-on-demand to be particularly significant in retail stores—but it's already a significant part of Amazon. "We already have many [print-on-demand books] in our catalog, but it's invisible to you, the customer. We use a number of companies that do the actual printing, but we mail them like regular books. They look like regular trade paperbacks."

Dyszel, Bill, "How to make an award-winning movie," *PC Magazine* 23:23 (December 28, 2004): 80-1.

I just had to include this charming story. Dyszel entered the 48 Hour Film Project's New York competition—a competition that gives teams two days, 48 hours, to prepare a finished short movie. Each film must include a compulsory character, a compulsory prop, and a specific line of dialogue; the genre for each film is decided by lot. Dyszel drew science fiction and always wanted to make a musical, so he made a 150-second sci-fi musical. Oh, and unlike most teams, he did it all himself, using consumer-level equipment and software. It's a fascinating read. You can even view the flick (http://go.pcmag.com/area2slash2), which picked up three awards.

Entlich, Richard, "FAQ: One last spin: Floppy disks head toward retirement," *RLG DigiNews* 8:8 (December 15, 2004). (Available at www.rlg.org)

As always *RLG DigiNews* offers valuable content that should be in THE LIBRARY STUFF. This issue includes a piece on *Ephemeral Cities*, a historical digital atlas project; the role of geographic location in metadata schemas and digital collections; and PREMIS—Preservation Metadata Implementation Strategies.

I'm only commenting on the latest in Richard Entlich's wonderful "FAQ" series. He begins with a question: "It appears the floppy disk is going the way of the long playing record and the rotary dial phone. Is there any cause for concern?" In three text-heavy pages, he explores that question.

Entlich notes that Macs haven't had diskette drives for a while now, that most notebooks don't include diskette drives (he calls them "floppy drives," although there hasn't been a true floppy for *years* now), and that Dell dropped diskette drives on high-end computers two years ago. My new work computer doesn't have a diskette drive; it turns out that the one on my old work computer didn't work very well anyway. (Lexar Jump Drive to the rescue!)

IBM introduced the 8" floppy in 1971; Shugart issued the first "minifloppy" (and last true floppy), the 5.25" form, in 1976. Sony created the 3.5" microdiskette in 1981; Apple adopted it for the Mac in 1984, and it became the dominant form over some years. There have been many other sizes and densities; most of them had no market impact. More history concerns physical stability, compatibility over the years, and why diskettes have lasted this long despite being fairly fragile and the slowest storage medium around. Finally, Entlich comes to the bottom line: "Ultimately, it is the floppy's limited storage capacity that has spelled its doom." It doesn't help that recent 3.5" drives have been so cheaply made that diskettes don't work reliably any more.

The discussion of salvaging data still on diskettes raises issues central to *RLG DigiNews*—and it's getting harder. I decided not to even try rescuing the data on three boxes of 8" diskettes; although they represented a year's worth of work, it was programming for a computer and operating system (Datapoint Databus) that probably no longer exist anyway. If I did want to rescue the files, it wouldn't be easy. For that matter, *MARC for Library Use* was written on 5.25" diskettes under CP/M; if I haven't migrated the files, it would probably be easier to scan and OCR the book than to try to retrieve the files. (I did a truly clever thing: Migrated all my old text files to Zip cartridges. And recycled the computer and Zip drive. Fortunately, I *think* I also migrated all the text files to CD-R and zipped archives on my hard disk as well. If not…)

Diskette drives are obsolescent now. "Whether it takes two years or five or ten, all floppy disk formats will be obsolete within the foreseeable future"—not obsolescent, but obsolete. Any data on them now

"should be considered endangered" for reasons explained in the FAQ.

Entlich draws lessons from the history of diskettes. "All media is subject to obsolescence. Media types that are less mainstream, less standardized, and less widely adopted are hit harder and faster by obsolescence. If one is paying any attention, there is generally ample opportunity to recognize that a medium is headed for obsolescence at a time when migration can be accomplished relatively painlessly and inexpensively." Entlich doesn't see a single obvious replacement for diskettes. I think he's right—neither CD-R/recordable DVD nor USB flash drives have all the desirable characteristics of diskettes, although both offer more useful capacities.

An excellent, readable article, typical of Entlich's work. **Highly recommended**.

Guenther, Kim, "Pull up a chair and stay awhile: Strategies to maximize site stickiness," *Online* 28:6 (November/December 2004): 55-57.

"Stickiness is the average time a user spends on a site and the frequency of his or her visits to the site." I'm a little surprised by that definition; I would have thought it only meant time spent. To me, frequency is part of loyalty or worth—but that's a quibble.

"The idea of stickiness is based on the premise that the longer a user remains on the Web site, the more potential exists to influence that user's behavior." Maybe, if it's the right kind of stickiness—but one reason Google is so popular is that the clean, efficient screen gets you in, out, and off in a hurry.

Guenther gets that. As you read further in this **recommended** article (recommended for webmasters, at least), you learn to distinguish between good and bad stickiness and see some recommendations for a (good) sticky site. Three of those suggestions (keep it simple, deliver value with every "click," and don't waste my time) represent distinguishing factors between good and bad stickiness.

Library sites may tend to be like Google: You really only want people on long enough to do their business—but you want them to *like* using the site. Guenther's suggestions are worth considering.

Machrone, Bill, "Hack your gadgets," *PC Magazine* 23:18 (October 19, 2004): 74-75.

Maybe "good stuff" is the wrong section, but I had to mention this strange article about tricks to make your digital gadgets do more than was originally intended. That includes hacking a Canon EOS Digital Rebel to provide some firmware features of the more expensive EOS 10D, turning an Xbox into a Linux workstation, refilling inkjet cartridges, and turning Roombas into robots. It also includes getting around DVD limits—defeating region codes, for example.

Some of these may be interesting projects. Some may be exercises in futility (I'd put refilling inkjet cartridges in that category, particularly if you're using the new durable inks). And some, as Bill doesn't note, are almost certainly flat-out illegal—violating DMCA by subverting digital protection.

Notess, Greg R., "The changing information cycle," *Online* 28:5 (September/October 2004): 40-42.

Notess recognizes that "the information journey on the Internet differs from a similar search in bibliographic or full-text databases" and discusses some of those differences. "The Internet" here is shorthand for the open web, since most bibliographic and full-text databases are also provided via the internet.

Perhaps the most important lesson is that you need to go beyond a single page to be reasonably certain of your information. Even the most authoritative sites can have typographical errors; "it is so easy to post a Web page that much Web content fails to have significant editorial oversight." He recommends scanning *at least* the ten results that most search engines display as a default first page.

Notess finds "that I am working on retraining myself to dig more deeply on the Web, to look more broadly at the range of answers." Anyone who relies on the open web for factual information should consider that process. **Recommended**.

## Session Report: ALA Midwinter 2005
# ACRL Current Topics Discussion Group

Barbara Blummer,
Center for Computing Sciences

The Association for College and Research Libraries Current Topics Discussion Group included three excellent presentations at ALA Midwinter.

## Institutional repositories: Their place in scholarly communication

Michael Keller (University Librarian and Director of Academic Information Resources at Stanford University) explored various conditions fostering the development of institutional repositories. He noted the popularity of digital methods rather than print for publishing, teaching and communication in the academic community. He also outlined environmental factors outside of the university that boosted the growth of digital works including mass digitalization projects, the publication of government information in digital format, and major research libraries' development of open source methods and applications for digital repositories including: LOCKSS, DSpace, and Fedora. Additional factors supporting institutional repositories discussed by Keller included: Internet 2 networks, improvements in search/discovery (Google Scholar), and digital discovery and retrieval mechanisms in development.

The presentation also provided ideas for populating digital repositories which Keller referred to as "Communities and Content" and included categories such as students and teaching, faculty schools and departments, and alumni. The speaker believes libraries need various "Tools and Services" to facilitate the creation of institutional repositories such as metadata, data management, project maintenance, digitization services and digital enterprise. Moreover, he views institutional repositories as offering new roles for librarians in digital selection, intellectual access through indexing, preservation, search assistance, and support through extraction, analysis, and presentation. The talk concluded with a list of hurdles to creating and supporting institutional repositories such as authenticating digital objects, obtaining release permissions from creators, multiple versions and multiple localities, and new file and MIME types.

This was an informative discussion on the background and growth of institutional repositories in academic institutions. The list of material to be included in selection remains especially relevant as well as Keller's description of new roles for librarians in digital repositories and problems with their installation and maintenance. Unfortunately Keller neglected to relay information pertaining to Stanford's experience establishing their institutional repository, which would have underscored his expertise in delivering the talk.

## UThink: Blogs at the University of Minnesota

Shane Nackerud (webmaster, University of Minnesota) discussed his involvement in facilitating the library's hosting of blogs at the University, beginning with a white paper he authored in June 2003. According to Nackerud, the library's goals for sponsoring blogs included: promoting intellectual freedom, building communities of interest, enhancing traditional academic enterprise, retaining the cultural memory of the institution, and changing perceptions. Although September 2003 marked the official start of the project, blogging was slow to spread in the University. He credited the appearance of a description of the blog project on the library's web page with their eventual adoption throughout the campus by undergraduates, graduates, as well as faculty.

Nackerud's topic held the interest of the audience throughout his presentation. He noted the library's incorporation of additional technologies for blogs such as RSS feeds as well as SFX, which allows users to post citations on blogs. This was especially well received by listeners. His description of the University of Minnesota's publishing opportunities with "Into the Blogosphere" (a scholarly publication hosted by the library) also sparked the interest of the group. Nackerud's concluding remarks centered on the library's future plans for the UThink project including: an upgrade to Movable Type 3, creating archiving procedures and policies, expanding the project to other University of Minnesota campuses, building a more robust blog search engine and seeking out additional publishing opportunities. He also discussed his plans to foster community building through the organization of blogs by majors, departments and affiliations.

Nackerud peppered his talk with positive and negative experiences hosting blogs including the posting of offensive materials and the slow adoption of blogs. His inclusion of usage statistics, 900 blogs and 12,800 individual posts by April of 2004, provided an illustration of their popularity on campus. The speaker's description of the type of blogs created in the University, such as the legislative network blog which reports on legislative events affecting the campus, teaching blogs, and blogs from PhD candidates reporting their research, added new insight to the various possibilities for blogging. His goals, and especially his library's hope to lend legitimacy to blogs by hosting them, underscores the importance of this new communication medium for all institutions.

*Building an institutional repository*

Margaret Branschofsky (Massachusetts Institute of Technology Libraries (MIT) Libraries) provided an overview of institutional repositories from the MIT perspective. She characterized institutional repositories as institutionally defined, containing scholarly content, cumulative and perpetual, and providing interoperability on an open access platform. Her presentation centered on the development of DSpace, the institutional repository open source software used at MIT. According to the speaker, DSpace originated as a joint venture between MIT Libraries and Hewlett Packard who sought to provide space for the school's faculty, digital research, and teaching materials utilizing open source software. The project's developers envisioned content as including technical reports, working papers, conference papers, preprints, post prints, books, thesis, datasets, course materials, and digitalized library collections in a variety of formats such as text, images, audio and video.

The second half of Branschofsky's talk chronicled the project's successes and failures since its debut in November 2002. The speaker related how DSpace's developers granted MIT's faculty wide leeway in populating the database. MIT Libraries provided DSpace's contributors with tools such as a metadata template as well as final decision making on submission and access rights to the material. DSpace's only requirement, according to Branschofsky, was that every item be available to MIT staff. Moreover, DSpace also allows authors to retain copyright since the license only grants the institutional repository a non-exclusive right to acquire, manage, and distribute the material. Still, according to the speaker, MIT's staff was slow to adopt the institutional repository concept despite the libraries' mass marketing campaign which included press releases and presentations. MIT Libraries were ultimately forced to hire a marketing consultant who advised them to identify an audience, target the decision maker and promote the institutional repository project to that individual. According to Branschofsky, DSpace currently receives 2563 hits per day and includes 5699 items from 17 communities and 48 collections.

*Comments*

This was an excellent presentation that provided useful information for libraries interested in establishing an institutional repository for their communities. The speaker's detailed description on the organization of the repository in communities and sub-communities as well as list of possible content provided a graphic illustration of the database at MIT. Moreover, her talk also included strategies for success which centered on anticipating questions from faculty on topics such as the differences between institutional repositories and personal websites, intellectual property issues, digital presentation of materials and preservation. Her final advice to potential institutional repository librarians centered on identifying barriers, speaking to non-users, and testing new approaches.

## Session Report: ALA Midwinter 2005
# Electronic Resource Management Systems

Barbara Blummer,
Center for Computing Sciences

This meeting of the Association for Library Collections and Technical Services (ALCTS) Electronic Resources Interest Group included presentations by Tony Harvell from University of California San Diego (UCSD) Libraries and Ivy Anderson and Ellen Duranceau representing Harvard University Libraries and the Massachusetts Institute of Technology Libraries (MIT) respectively, on their collaborative efforts with integrated library system (ILS) vendors to develop an electronic resource management system (ERM).

*Implementing the Innovative Electronic Resources Management System*

Harvell's presentation traced the Libraries' installation of Innovative's ERM product in the fall of 2003. Harvell began with an overview of UCSD Libraries online journal collection, which includes more than 300 electronic resources packages representing nearly 10,000 licensed serials. UCSD Libraries, like many libraries, housed information pertaining to their ejournals in numerous places including paper files as well as their ILS system. The haphazard tracking of ejournals created problems for renewals, statistical reporting, and updating holdings information at UCSD Libraries.

Harvell described how Innovative's ERM system provided new record types to match the needs of electronic journals such as resource, license, and order records. The system also displays links among record

types, including holdings records, to describe relationships. In addition, the software contains a tickler which allows advance email notification of events such as renewals. Moreover, it provides contact information as well as incident logs to track performance problems. Two especially interesting features include the ability to display license terms in the OPAC at the title level and the customization capabilities of the public interface to display information from the resource record. Harvell also described how holdings information can be updated through batch processes.

The speaker concluded his presentation with pointers on a successful implementation of an ERM system such as defining what to track and determining how to track it as well as establishing priorities and training staff. Harvell believes it is especially important to develop local standards and involve all staff including public services in creating the system. His slides on "What Works" (including better integration of electronic resources in the ILS) and "What Still Isn't Working" (populating resource and licensing records is slow) provides valuable information to libraries considering purchasing the product.

### From Greenbox to Verde

Duranceau and Anderson discussed Harvard and MIT's collaboration with Ex Libris to create a viable ERM system, Verde. This talk, which provided a theoretical perspective on ERM creation rather than a practical ERM application, began with a timeline tracing significant events in Verde's development. The project commenced in December 2002 with the two librarians defining the functionality of the fields and data elements. By March 2004 the pair had created a data model and an agreement between the two institutions sealed the development of the project. Duranceau and Anderson credited the guidelines for functional requirements established by the Digital Library Federation Electronic Resource Management Initiative (DLF ERMI) as fostering the creation of the data model.

The speakers outlined some key features of the Verde software such as a knowledgebase approach, no public end user interface and the ability to integrate with Ex Libris Aleph Acquisitions Module. The speakers remained especially excited about Verde's knowledgebase approach which allows SFX or MetaLib's central knowledgebase to prepopulate fields. Libraries shouldn't hesitate over Verde's lack of a public inter-

face since the software can deliver information through Ex Libris's OPAC, MetaLib, or any third party application via web services. According to Duranceau and Anderson the product is designed to accommodate consortia, but they did not elaborate. Their presentation also included a demonstration of Verde which included searching for an ejournal, and editing the acquisitions and license records. The speakers concluded with a statement that adherence to DLF ERMI specifications should facilitate interoperability and portability.

### Comments

Each presentation complemented the other and provided the audience with an overview of the development and utilization of ERM systems in academic libraries. Harvell's talk focused on the practical implementation and utilization of an ILS ERM product with his outline of record types and hints for successful applications as well as what the system can and can't do. Duranceau and Anderson's presentation, on the other hand, discussed the creation of the software which included utilization of data elements and a data model to develop a viable system based on standards and guidelines from the DLF ERMI.

Barbara Blummer, bablumm@super.org, is reference librarian at the Center for Computing Sciences.

# Masthead

URL: cites.boisestate.edu/civ5i3.pdf