

Cites & Insights

Crawford at Large

Libraries • Policy • Technology • Media

Sponsored by YBP Library Services

Volume 7, Number 3: March 2007

ISSN 1534-0937

Walt Crawford

Bibs & Blather

Who's Out There?

It's been 21 months since the last "Readership Patterns" discussion (*C&I* 5:8, June 2005). Since then, *Cites & Insights* has changed hosts and included two or three major special pieces. Time for another check.

First the baseline: My expectations and hopes for readership. I expect to have at least a few hundred readers; otherwise, this might not be worth doing. I *hope* most essays will eventually reach 1,500 readers or more, although that's not critical. Anything over 2,000 is gravy: I'm delighted, to be sure.

When I looked at 2003 and 2004 readership in early 2005, I found most issues in 2003 (volume 3) had between 1,300 and 2,500 downloads, with one showing 4,500 downloads. In 2004, half the issues exceeded 2,400 downloads (three over 3,000) and four more had 1,800 to 2,400 downloads—but three others had fewer than 1,600 downloads, with one showing 1,100 (but that omitted a few hundred readers at a temporary site). Back then, there were no HTML separates, which made numbers easier but also made it harder to spot essays of special interest.

Looking at patterns through February 1, 2007, several things are clear:

- Readership continues to grow. At the old website (cites.boisestate.edu), the first half of 2003 showed 2,200 to 4,700 visitors per month (4 months under 4,000); the second half, 4,460 to 6,730 (4 months under 5,000). The first half of 2004 showed 4,800 to 6,806 visitors per month (4 months between 5,000 and 6,000); the second half, 6,113 to 7,420 (4 months under 7,000). Come 2005, the first half ran anywhere from 7,560 to 12,291 (3 months over 11,400), the second half 9,405 to 15,634 (5 months over 12,800). The first

half of 2006—noting that *C&I* moved in July 2006—included the first and only month over 20,000 (January 2006, which included the Midwinter "Library 2.0" issue) and had three other months over 18,800—with none lower than 15,436.

- Here's where it gets interesting: The *last* six months of 2006 showed 14,600 to 18,774 visitors per month, and while the highest number was July, there were only two months under 15,000. That's with *no new issues* and all of my posts and other publicity (and the home page) pointing to citesandinsights.info. There's a *lot* of residual readership coming from other sources. January 2007 isn't bad either: 14,674 visitors.

Inside This Issue

Old Media/New Media.....	3
Trends & Quick Takes	5
Net Media Perspective: Wikipedia Revisited	12
My Back Pages.....	21

- I don't have monthly readership for the new site, but over the 6.5 months it's been operating there have been 46,000 sessions—roughly an average of 7,000 per month. That's a substantial drop but it's also a different log analysis system. I trust indirect traffic will rise over time—but if it doesn't, that's OK.
- Since *C&I* moved to cites.boisestate.edu in December 2002, more than 133,000 unique IP addresses have visited the site—a number that simply astonishes me. The number for the new site: A little over 18,000.

For individual issues and essays, I've added the two PDF figures—but haven't attempted to integrate HTML counts with PDF downloads. What I see is that readership is spread out over several years, which is surprising but gratifying. Given the patterns I see, I

won't draw conclusions about the relatively low download figures for issues 6:11 through 6:14 (all between 1,200 and 1,500, *not* including HTML separates)—it's too early to say much.

The numbers here are somewhat startling. Here they are, volume by volume, noting that Volume 1 and Volume 2 (and the inaugural issue) originally appeared on my personal AT&T account, so these numbers are really *all* delayed readership:

- Six of the issues in Volume 1 have more than 2,000 downloads since moving to boisesstate.edu; the other seven all exceed 1,200.
- Four of the issues in Volume 2 have more than 2,000; the others all exceed 1,200.
- One issue in Volume 3 (the CIPA special) now shows more than 10,000 downloads. Six others exceed 5,000, six more exceed 3,000—and one had 2,638. That was issue 3:2 (February 2003), with ten different sections in 18 pages. You don't read *C&I* for “blog post equivalents” or “short column equivalents.”
- Volume 4 shows five issues over 5,000, three more over 4,000, five more over 3,000—and one with 2,356. That's 4:10, the one first published at AT&T.net due to server problems; it's fair to estimate that the count is at least 500 to 600 low.
- Volume 5 has two issues over 4,000, six over 3,000, and the remainder all over 2,000.
- Volume 6 has one issue over 10,000 (guess which?), two over 3,000, six over 2,000, and the remainder over 1,200. The only one of the remainder that's more than five months old is issue 6:5, the Diamond Anniversary issue—and I think I knew 75 mini-essays weren't a great idea as soon as I published it.

If someone asked what my readership was, I'd say “at least 1,500 to 2,000 regular readers (including HTML and passalong PDF), with more over time.” But, of course, I factor in some portion of HTML downloads for recent issues. What about those numbers?

- The monster, as you'd expect, is v6i2a, LIBRARY 2.0 AND “LIBRARY 2.0”—12,538 visits to the HTML page *along with* 11,555 PDF downloads. That comes out to over 24,000—about 40 times the readership I originally hoped to attract here and about 10 times what I consider enormous success.
- Only one other essay had more than 10,000 HTML hits: v5i10b, You can probably guess

that one too: “Investigating the Biblioblogosphere.” Plus 4,795 PDF downloads...

- Other than internal pages (oldvol, about, citoc, cifaq), HTML hits drop fast after that. v5i13a, “Life Trumps Blogging,” added 5,333 hits to a 2,600-download issue; v5i7d, “Weblogging Ethics and Impact,” added 3,016 hits to a 3,855-download issue. All the others are under 3,000, with 19 over 2,000 (all but two of them from volumes 5 and 6) and at least 50 more over 1,000. If I had to estimate, I'd say most substantive essays eventually pick up more than 1,000 HTML readers.

That's it—too much, I'm sure. Incidentally, at least 150 different nations show up with three or more visitors at cites.boisestate.edu; 71 have 100 or more and 22 have 1,000 or more. Top 10, in descending order: United States, UK, Canada, Germany, China, France, Australia, Norway, Netherlands, Japan.

Language Grump: Lose and Loose

This one's just dumb, but it sure is common: “loose” as a verb and “lose” as an adjective. That error passes any spellchecker and if people pay attention to Word's grammar check, it probably doesn't spot the problem. (Yes, I know “loose” *can* be a verb—but any time I've seen it used that way, they meant “lose.” Few people say things like “Loose my heart from these binds.”) Whether it's a spelling error or a grammatical error, it's unfortunate.

I can't imagine any *C&I* reader doesn't know the difference, but just in case, consider antonyms. Lose is the opposite of find; loose is the opposite of tight. You can't tight an object, and you can't have a lose command of the English language.

Another Peculiar Issue

The first two issues of Volume 7 were PERSPECTIVE-heavy, since FINDING A BALANCE is a category of PERSPECTIVE. This issue's more random, although it appears “media” is the (accidental) theme for most of it. Meanwhile, it's been a while since a piece on LIBRARY ACCESS TO SCHOLARSHIP—and longer since a copyright essay. I'm planning an LAtS piece for next issue; copyright-related issues are light at the moment.

The real reason for the peculiar nature of this issue? Simple.

Remember BIBS & BLATHER in *C&I* 7:1—specifically, “This Year's Plans”? The part about book ideas?

It's now pretty nearly certain that I'll self-publish the first book in a possible "C&I Books" series. The rough draft is done. As I write this, I'm a third of the way through revisions. I still need to design the book template, do the rest of the revisions, add index entries (oh, joy) and do the final publishing bits (cover design, Lulu storefront, upload and publish). It's possible the book will be available before the April C&I is out or appear simultaneously with that issue. Unless something goes wrong, it's *likely* it will be out before the May issue.

Some of the energy that might otherwise go into *Cites & Insights* has been diverted to the book as it's begun to seem more plausible. I'm not apologizing—I think the first two issues of the year were both strong and this one isn't bad—but there it is.

There's crossover. You've seen a rough draft of Chapter 2 and a not-quite-final draft of Chapter 15. A third chapter may appear in the next issue. As soon as the book's complete I'll describe it here.

Check *Walt at Random* for updates. If you're not getting posts, unsubscribe and resubscribe with the first browser-icon option (a Feedburner feed).

Old Media/New Media

Facing the Music

Will legal music downloads become real purchases—that is, with fair use and first-sale rights and without Digital Restrictions Management? Some of them already are (EMusic, for example) and it's possible they'll all go that way.

If that happens, I'm sure Steve Jobs will claim and get most of the credit. His recent call for the big labels to drop DRM is somewhat hypocritical, given that iTunes won't sell DRM-free downloads from independent labels that *don't want* DRM (and sell that way through EMusic) and that he absolutely refuses to level the playing field by licensing iTunes' DRM to other stores and players. One of his points is pretty much on the money: Better than 90% of all music legally sold today is *already* DRM-free—because CDs outsell downloads by at least ten to one.

Eliot Van Buskirk offered seven reasons major labels might prefer MP3 (without DRM) in a January 8, 2007 *Wired News* column—*before* Jobs' speech. His first reason may be right but the argument behind it is *Wired* hyperbole. The reason: "The labels don't have a choice"—which boils down to them not being willing

to let Steve Jobs monopolize distribution. That's preceded by "When CD sales finally tank completely," but there's no real evidence that everyone wants to give up physical carriers (I sure don't). If you're *Wired*, digital is always better and everybody agrees, so it's *inevitable*. Most clear-headed discussions of the drop in CD sales suggest that labels churning out mountains of repetitive crap has more to do with it than either legal downloads or a flight from CDs—except to the extent that buying single tracks lets you filter out most of the crap. (That section of Buskirk's article notes EMusic is the second largest online music store "because it sells digital music rather than digital rights." I know it's the one I'd use right now if I was ready to buy downloads.)

His other reasons? Apple might be forced into interoperability (there are class action lawsuits asserting antitrust behavior). Thomson, which licenses one MP3 algorithm, is suggesting watermarked MP3s; I'm not sure that means much. There are rumors that Amazon plans to sell MP3s, which might mean more serious competition. One Sony *Electronics* official said DRM will become less important—but there's a disconnect between Sony's hardware side and Sony's music-publishing side. "People love AllofMP3.com" (a Russian outfit that sells cheap and "questionably legal" MP3 downloads—but there's really no way to *pay* for those downloads anymore). MP3 has future options—you can improve quality (already easy to do by lowering compression) and add surround sound. I'm not sure these all play out, but it would be nice to see part of Big Media begin to accept that DRM is more trouble than it's worth.

Speaking of sound recordings, I ran into a silly piece from the January 14, 2007 *Houston Press*, "DISConnect" by John Nova Lomax. It's a long piece, five pages of small single-spaced type. The tease: "Sales of CDs are falling faster than you can say iPod." That's nonsense—there's continued erosion, but it's *not* "faster and steeper every year" (as the article says). I'm curious as to why an EMI hotshot would say "The CD as it is right now is dead." It's certainly the bulk of EMI's music revenue.

How bad is this article? It says *worldwide* CD sales are \$6.45 billion, presumably meaning for the year 2005, since 2006 revenues wouldn't be known that early—but *U.S.* CD sales alone were \$10.5 billion in 2005, according to the RIAA, and it's hard to believe that international sales amounted to "minus \$4 billion." It is true that the first half of 2006 was down

sharply from the first half of 2005. It's also true that 2004 sales were *higher* than 2003; "faster and steeper every year" is just wrong except on a one-year basis. The story also says \$945 million worldwide for digital sales, or about one-seventh of CD sales—but U.S. 2005 figures were about \$500 million in downloads, or about one-twentieth of CD sales, and even the first half of 2006 was \$400 million downloads to \$4.06 billion CDs, or about one-tenth of CD sales.

Lomax flatly says that CD sound isn't better than vinyl and "most audiophiles argue that their sound is inferior," which strikes me as a considerable overstatement. Later in the story, Lomax quotes a local musician as to *why* vinyl is better: it's *the crackle* that makes it music! Lomax tells us the CD "really will be dead" within five years. I find that unlikely. As for the resurgence of vinyl, that one is amusing: According to RIAA, vinyl LPs sold \$14.2 million in 2005 (about 0.14% of CD sales), and the combination of vinyl LPs, cassettes and CD singles amounted to \$12.4 million in the first half of 2006 [the numbers are so small that RIAA doesn't break them out for the first-half tables]. It does make you wonder about \$90,000 turntables: Is it possible that most of vinyl's "resurgence" is turntables, cartridges and preamps for a few thousand wealthy advocates, not LPs themselves?

There is some recognition of what's really going on. "We really *are* awash in a sea of crap these days." Maybe it's because there's too much music, it's too easy to get and it doesn't mean much to people as a result. Maybe it's because "in any society, there are going to be a few truly talented musicians"—but almost anyone can get their music out these days, and that may not be a good thing. Certainly the mind-numbing playlists at commercial radio stations (and the dominance of commercial radio by a handful of companies) don't help.

A strange article, mixing bad numbers, lots of nostalgia for the glory days of LPs, claims that we can't concentrate long enough for whole songs or albums these days (quoting an absurd McLuhan maxim, "The future of the book is the blurb"—and the future of McLuhan was apparently sloganeering), a misreading of EMusic as being entirely a subscription service and claims that vinyl has "totally made a comeback" and "just sounds better than CD."

Video Killed the TV Star?

Maybe not. YouTube and its competitors serve a variety of interesting, if mostly silly, purposes—but are

you really going to replace NBC, ABC, CBS, Fox, CW, HBO, Showtime, Comedy Central, PBS and the rest on your big-screen HDTV with the stuff and the video quality you get on YouTube?

Really?

Nope, didn't think so. If you've tried expanding a typical YouTube or Google Video clip to full screen, you know how painful it can be.

Maybe you'll go the other way—watch your TV shows on your mobile phone. *Sure* you will. When London researchers studied preferences of people with video-capable cell phones (as reported in *New Media*), they found people wanted to watch full-length programs "on their TVs at home." As a researcher put it, "there aren't enough viewing occasions on the go for long shows." It's also a matter of quality: "What looks great on widescreen TV looks disastrous on mobile. With sport, you can't see the ball." Mobile video content has to suit the small screen: not much detail, not much going on, not too long. Two to three minutes seems about right—and, supposedly, people will pay \$1.70 or so for a good two or three minute clip. The strangest item in the study: About 40% of viewing video on a cell phone was *at home*.

Sean Captain wrote "Forget YouTube," which appeared May 8, 2006 at *Slate*. He's not saying YouTube is doomed. The key is the subhead: "Your laptop will never replace your TV." By "never" he means "not for the next decade at least." Partly that's because the internet is a lousy high-def video distribution network: There's just not enough bandwidth and packet transmission isn't the way to assure smooth high-def video. "Today's Web video looks rather good because we see it on very small screens"—but most of it doesn't even reach VHS quality. "ABC's 700-by-394 pixel videos may look decent in a small window on a laptop but will likely be unwatchable on a 1280-by-768 pixel, 50-inch plasma TV." The hallmark these days is a 1920x1080 pixel screen, to be sure. I'll admit I was surprised at how good one TV show (we'd missed an episode) actually looked on my 19" Sony LCD monitor—but it wasn't VHS quality and would have been barely adequate on our 32" [SD]TV.

Clay Shirky doesn't see it that way. In a January 3, 2007 *Many2Many* post, he asserts that Mark Cuban (a successful impresario of HD broadcasting) "doesn't understand television" when he says that internet video isn't going to replace HDTV. Cuban overstated the reverse case—or maybe he didn't. "HDTV is the Internet video killer" may be right in terms of any-

thing more than short videos although it's wrong in terms of doing away with web video entirely. Shirky draws the analogy of audiophiles saying "MP3s won't catch on"—curiously, by quoting a paragraph that says nothing of the sort. (The writer noted that MP3 sound quality for downloads was still radio quality, and until that changes he wasn't all that interested.) Shirky's snarky analysis of that non-statement seems to assume that MP3s have taken over—even though CDs still outsell them more than ten to one.

For Shirky, it's simple. Since the form of a video and its method of delivery *can* be separated, "they will be, because users prefer it that way." We all think like Clay Shirky. And, of course, "People like to watch, but they also like to create, and to share." True, although most people who watch don't want to create—and it's also true that *there's no conflict between the two*. Do people who create three-minute videos on YouTube stop watching *Lost* and *Desperate Housewives* and *Bones*? Why would they?

Shirky seems to think people don't give a damn about video quality, that people "don't optimize, they satisfice." Apparently YouTube is the "proof" that internet video is good enough. His conclusion: "YouTube is the HDTV killer. Count on it." I'm amazed Shirky is so confident he knows more about HDTV than someone who actually runs HDTV services. I guess theory beats practice every time.

The comment stream was interesting. Mark Cuban was first, noting that "people will take the path of least resistance. They always do." Today, getting HD programming may be a slight nuisance—but that will change. Meanwhile, even *getting* internet video to the TV is a much bigger nuisance. Most of us don't want our PCs to be in our living rooms. Cuban notes something I suspect is true: Most internet video is watched during office hours. Cuban also notes that people creating their own music hasn't resulted in a profusion of *successful* sites; by and large, "no one cares."

Another commenter notes that, "everything being equal—higher fidelity always wins," but everything isn't equal. Another makes the obvious point that it isn't either/or, it's both: HDTV really isn't going away (Shirky's wrong), and Internet video isn't completely going away (Cuban's wrong). One commenter claims it's been "proven in spades" that "quantity of choice beats quality by a mile" and that breaking TV shows down into little pieces makes them "so much more compelling." That commenter seems to think that

"people" (most? all?) already prefer web video to TV. I don't think the numbers bear that out or are likely to.

Media Life for February 8, 2007 does the unthinkable in this regard: It looks at research studies (in a story by Kevin Downey). Two studies considered here show that "the online video craze isn't nearly as big as it sometimes seems. And more significant to advertisers, it's having no effect on TV viewing." The reporter gets it right: "online video will be yet another media option." Online video is "about a minority of people doing it a lot"—maybe 4% of adults spending an hour a day with online video, compared with 93% of adults watching TV at least an hour a day. "That is not mass at this point. Usage is increasing but the user base isn't increasing much." That sounds right. I might click on one video link in a blog, but I won't go clicking beyond that—after about seven minutes, it gets old. I can see that some people would find it fascinating and continue exploring for quite a while—quite possibly while they're supposed to be doing some white-collar job. The *Media Life* story (note that *Media Life* is new media—an emagazine) gets it right about new media "versus" old media in general:

The concerns that online video will negatively affect TV viewing are reminiscent of fears raised over the past few decades about each new medium that came along.

It's almost always the case that consumers simply squeeze in time for new media in addition to their usage of existing media.

Or, if Mark Cuban is right about web video, it requires "multitasking" at work rather than cutting down on home TV consumption.

Oh, about the disastrous state of daily newspapers? Turns out the numbers aren't quite what they seem—but that's a discussion for another issue.

Trends & Quick Takes

The Snark Factor

Some readers may wonder why something winds up in TRENDS & QUICK TAKES rather than MY BACK PAGES or vice-versa. For both of you, the heading above should be a tipoff. MY BACK PAGES is *mostly* snark, as is the final feature in so many magazines. It's stuff I find impossible to take seriously or that's a little too revealing for its own good. Snark certainly appears elsewhere, but it's rarely the *primary* reason for discussing something.

At one point, I believe TRENDS & QUICK TAKES was mostly about trendspotting. That point is some-

where in the past. Now it's a catchall for brief essays that don't fit into one of the current categories (or that fits into a semi-dormant category like THE GOOD STUFF). Stuff that appears here should be at least marginally worth thinking about. If I argue strenuously against something, it's because I think the argument matters. Otherwise, it would go in MY BACK PAGES.

Now that *C&I* has passed the Crawford Test (Tom Wilson's name for my informal measure that an ejournal lasting at least six years can be considered a "lasting title"), I don't consider *Cites & Insights* experimental (as I said while back). That doesn't mean it's not continually changing or that I'm not making it up as I go along. Any good magazine or journal changes over time. Why should this peculiar publication be any different?

Speaking of peculiar publications, I should once again salute *Ex Libris* and Marylaine Block. When *C&I* began, there were four more-or-less comparable publications: *Ex Libris* (a weekly consisting of one relatively brief essay), *Library Juice* (also weekly but typically including multiple segments), *NewBreed Librarian* (bimonthly, with several essays in each issue) and *FOS Newsletter* (monthly, with several essays). All were labors of love. All published worthwhile material. All were free online. *C&I*, uniquely, was designed to be printed.

NewBreed Librarian's gone. *Library Juice* is a blog and a book publisher but the ejournal's gone. *FOS Newsletter* became *SPARC Open Access Newsletter* and is now formally part of SPARC. Two independents remain: *Ex Libris* (not always weekly but still going strong) and *Cites & Insights* (which has modest sponsorship that comes without editorial control). I trust *Ex Libris* works to increase Marylaine Block's profile. I know *Cites & Insights* makes a difference, sometimes a big difference. I can't suggest that anyone else should follow in either of our footsteps. Blogs are a lot easier and formal publications look better on your vita.

Finding the Good User-Generated Stuff

Jon Pareles wrote "2006, brought to you by you" in the Music section of the *New York Times* on December 10, 2006; Nicholas Carr commented on it in "Lost in the shitstream" at *Rough type* on the same day. Both discuss user-generated content—all those blogs, most of what's on YouTube, the "homemade art" on MySpace (and many other sites).

Pareles prefers "self-expression" to "user-generated content" and notes:

It's homemade art independently distributed and inventively promoted. It's borrowed art that has been warped, wrecked, mocked and sometimes improved... It's word of mouth that can reach the entire world.

It's often inept, but every so often it's inspired, or at least worth a mouse click...

He cites two "stars" from web self-expression: "video diarist" Lonelygirl15 (a fictional creation, really an actress being paid lousy wages) and a power-pop band "whose treadmill choreography earned far more plays than its albums." It's free labor—people supply all that raw material for nothing. But what do you do with it all?

In utopian terms the great abundance of self-expression puts an end to the old, supposedly wrongheaded gate-keeping mechanisms: hit-driven recording companies, hidebound movie studios, timid broadcast radio stations, trend-seeking media coverage. But toss out those old obstacles to creativity and, lo and behold, people begin to crave a new set of filters.

Pareles mentions that user-generated content isn't exactly new. Remember "America's Funniest Home Video"? Now, thanks to the web, it *can* reach wider audiences more easily and doesn't have to be ludicrous to "succeed." The article discusses ways some "traditional" creators are leveraging new tools—encouraging users to remix their songs, for example. He also notes that culture has never been monolithic and that contemporary fragmentation only goes so far. And there *are* various kinds of new filters.

The open question is whether those new, quirky, homemade filters will find better art than the old, crassly commercial ones. The most-played songs from unsigned bands on MySpace...tend to be as sappy as anything on the radio; the most-viewed videos on YouTube are novelty bits, and proudly dorky...

The promise of all the self-expression online is that genius will reach the public with fewer obstacles, bypassing the entrenched media. The reality is that genius has a bigger junk pile to climb out of than ever, one that requires just as much hustle and ingenuity as the old distribution system.

Carr's bias can be guessed from the s-word in his post title. It's a brief post raising questions about the "demand side" for culture:

As the flood of free, immediately and universally accessible user-generated and -filtered content grows, will the audience for well-crafted work shrink? Will we all readjust our tastes and expectations to the easy pleasures of the shitstream? I don't mean to sound an overly baleful note here. It would be a mistake (thank goodness) to think the motivations of the artist and the craftsman can be reduced to a set of signals from the marketplace. But

it would also be a mistake to think those motivations exist outside the influence of those signals. Even in the sphere of culture, demand drives supply.

By my lights, Carr is way too baleful, both because traditional media isn't disappearing and because there are all sorts of filters. My immediate response was to jot down "Pandora"—and that's a good response. Pandora is *not* a user filter. It does not rely on the supposed wisdom of the crowd. It does *work*, extremely well in my experience. So, for that matter, do some "wisdom of the crowd" filters. I don't believe Netflix' collaborative recommendations dumb down movie selections; if anything, quite the opposite.

Of course demand drives supply *to some extent*—but with lower-friction alternative distribution systems, niche suppliers can find niche demand more readily than in the past. This, I believe, is a good thing. That belief that tempers my impatience and occasional headshaking over the vast quantities of garbage on the web.

Comments on Carr's post are at least as interesting as the post itself. John Baschab was first up, noting that as he gets older, "my scarcest resource (by far) is time," so that brands and filters begin to matter more. "When I was a teenager I had what seemed to be endless amounts of time to discuss stereos, music, cars and other such pursuits. Not so now, and so brand-as-crutch is the order of the day."

Seth Finkelstein points out that, for text resources at least and in the larger web arena, there are new gatekeepers (filters)—and they work pretty much the same as the old ones did and do. "Captswing" makes the excellent point that Sturgeon's Law was with us even before the Gutenberg Bible—and that, historically, wider access to information and lowered barriers to entry for publication have been a total benefit for mankind. That, I believe, is a good and true point.

Maybe the bottom line is that old media aren't going away. Old media continue to filter in the ways that old media filter, absolutely blocking some forms of genius while promoting both good content and pablum that fits a formula. A variety of filtering techniques for new media are emerging. Those techniques seem no more likely to block true innovation and, because specialized audiences can yield specialized filters, somewhat more likely to encourage diverse creation. But then, I've always been a Candide at heart—and I know I could never do something like *Cites & Insights*, reaching 1,200 to 24,000 people, via traditional media.

Remix Culture?

Barbara Fister's "Money doesn't talk—it silences" (December 20, 2006 on *ACRLog*) relates to "Finding the good stuff..." above, but only indirectly. She's concerned about a new venture reported by the *Wall Street Journal*. The new venture, *Attributor*, takes your text and scans the web to see whether your stuff has been copied—so you can charge for the use or demand that it be taken down.

Quoting from the WSJ piece:

Attributor appears to go further than existing techniques for weeding out unauthorized uses of content online . . . [Company execs] claim to have cracked the thorny computer-science problem of scouring the entire Web by using undisclosed technology to efficiently process and comb through chunks of content. The company says it will have over 10 billion Web pages in its index before the end of this month

Fister's comments (excerpted):

So if I understand this, they copy web pages to see if they've been copied. And this kind of indexing, unlike the Google library project, doesn't violate anything because media companies might make money from it—and the heck with innocent bystanders whose work is copied into this massive database without permission. They aren't in business, so their rights don't matter. Right?

What corporations don't seem to understand is that a lot of the people involved in remix culture aren't interested in monetizing "intellectual property"...

The secret of Web 2.0 is that it revels in creation without worrying about artificially limiting access by charging a toll. This outpouring of creativity challenges the standard wisdom that the only incentive for creators is cash...

When will big media catch up to the idea that someone posting thirty seconds of Jon Stewart on YouTube isn't in it for the money—but drives audience to the show? If they could figure out that we're not all into monetizing, and stop spending so much money trying to make us stop, they could relax and reap the benefits of fans growing their market....

Some of us don't want to maximize revenue—we just wanna have fun.

This is a little tricky, as is "remix culture" in general. There's nothing wrong with me saying *C&I* has a BY-NC license, which means you can take it, "remix" it, whatever, as long as I get credit and you don't charge for it. *Attributor* would be of no use to me. As for a 30-second clip of Jon Stewart, that should count as fair use. But saying "we're not all into monetizing" doesn't automatically mean you get to "remix" any-

thing that suits your fancy. There's plenty of suitably licensed material for those who "wanna have fun," and *nobody's* stopping actual creators from not monetizing their work. You do *not* get to say how other creators should behave. You do *not* get to say "I want to reuse your creation without payment even though you don't want me to, and it's OK because I'm just having fun." That's not remixing, it's expropriation. (Yes, I'm a copyright centrist who believes creators have rights but not unlimited rights; hasn't that been obvious to all but control freaks?)

Placeblogging

I first heard of this *term* in a January 3, 2007 post with that title by George Needham at *It's all good*. There's a new site, Placeblogger (www.placeblogger.com), edited by Lisa Williams and "brought to you by the Center for Citizen Media, Pressthink, and H2otown." *H2otown* is one of the first placeblogs, coming from Watertown, MA. Here's part of how Placeblogger defines placeblogs:

Placeblogs [are] about the lived experience of a place. That experience may be news, or it may simply be about that part of our lives that isn't news but creates the texture of our daily lives: our commute, where we eat, conversations with our neighbors, the irritations and delights of living in a particular place among particular people...

Placeblogs spring from a fiercely non-generic America that's not about big-box retailers or the type of polarizing discussion about politics, culture, and the economy that's the product of journalism that happens at the 30,000 foot level. Often, they are a delightful and vivid look at cities, towns, and neighborhoods from an insider's point of view...

The site's home page is an odd mix of posts about local weblogs, headlines *from* such blogs, featured blogs and other stuff; I found it bemusing. There's a Google Group for placebloggers. The most useful aspect of the site, though, is probably its directory of place blogs (the site uses both "place blogs" and "placeblogs"). There's a feed that provides the first 200 characters of text from place blogs, which would probably be a very busy feed; the FAQ notes the reason for the 200-character (roughly 35-word) limitation: "we are serious about sending visitors to local sites, not keeping them here at Placeblogger."

As of February 19, 2007, the directory included blogs from 197 countries, including 4 from Niue, 7 from Brazil, 82 from Canada—and 1,646 from the U.S. (Originally, it was limited to the U.S.; no doubt some listings from some countries, particularly Eng-

lish-speaking countries, will grow rapidly as word of the site spreads. Australia has 17; the UK, 25. I'd be surprised if those weren't both triple-digit numbers by year's end.) Digging deeper, U.S. blogs organized by state include, for example, nine in Alaska, 27 in Ohio, 64 in Washington, 88 in Texas, 95 in Florida, 111 in New York—and 161 in California.

The state listing isn't quite alphabetical (it's probably alphabetic by two-letter code but displays as names), but every state's represented; the fewest blogs are in South Dakota (two) and Nebraska and Wyoming (six each). It's an inclusive list; Las Vegas and Reno each show five blogs, and in Reno's case I'd suggest that two or maybe three really qualify.

It's an interesting concept. Should a true placeblog be multiauthor (two of Reno's are personal blogs)? Should the term encompass local wikis as well (the subtitle at Placeblogger suggests that it should: "towards an annotated world with blogs, wikis, forums, maps...")?

Needham's take:

Wouldn't it make sense for libraries to be prominent in such placeblogs, or maybe even to start one if it doesn't already exist? Imagine how much the library staff could learn about its community by participating actively in this!

I'm inclined to agree—and wonder whether library blogs could be considered placeblogs in some cases. Of course, each placeblog is a distinct entity, and there may be placeblogs that have the wrong "feel" for the local library (e.g., a placeblog run by realtors might not only feel wrong, it might be out of the question), but as the directory and concept spread, it's worth considering for almost any public library.

Mark Lindner comments:

For once, a potential online community that seems like we should be in it from the start. Certainly public libraries, but even academics, especially state institutions, should be involved.

I'm not always the quickest guy, but I don't immediately see any downside to libraries being in this space.

Eric Hellman comments on a top-ranked place blogs (Baristanet), which covers his own town, and offers notes based on reading that blog since its inception:

A placeblog fills the same role as have "local rags" and "corner stores." Think of a placeblog as a police blotter with innuendo, gossip and an underpinning of real reporting... The weekly "open thread" is quite popular, and there are regular commenters who are really quite rude. Baristanet is always up-to-the minute when there are power outages, trees falling, houses burning, bears roaming, police cars chasing, houses selling, develop-

ments proposing, Sopranos filming or locals (such as Yogi Berra, Stephen Colbert) or ex-locals (Tom Cruise) making news.

Baristanet is financially viable because it provides a very focused audience for local advertisers. I doubt anyone is making a huge amount of money on Baristanet, but I bet it is attractive compared to freelancing...

It seems to me that a vital role for libraries is to start thinking about the best way to archive these placeblogs, just as they play a vital role in preserving local newspapers. Baristanet is a vibrant portrait of the life of my community in a way that my local newspaper has ceased to be.

I think this is a trend worth watching and maybe participating in. Hellman's final paragraph is worth thinking about in terms of your library's role in local history. I'm not ready to start a Mountain View placeblog—but then, we still *have* a local weekly, at least.

Marylaine Block wrote an excellent essay on "The Library's Place in Place Blogging." It's the February 16, 2007 edition of *Ex Libris*, #297. **Go read it.**

Maybe More than the Long Tail Deserves

There was a flurry of activity last summer connected to Chris Anderson's "Long Tail" stuff—from his nonsensical claim that "the era of the blockbuster is so over" and typical *Wired* overgeneralization to a batch of reviews and comments once the inevitable book actually appeared. Since libraries, magazines and book publishers have been in the "long tail" business for many decades (without needing to give it a hip name!), I won't spend too much space here.

A *New Yorker* review by John Cassidy lauds Anderson but notes that Alvin Toffler said much the same thing in 1980 (and was wrong, saying "no more mass production...no more mass entertainment") and points out that in this "post-blockbuster" era, seven of the ten all-time top-grossing movies came out since 2000, as have four of the top-selling novels. The reviewer concludes (correctly, I believe) that blockbusters and niche products will continue to coexist (as they *always* have, at least since blockbusters were possible)—and notes Anderson's blind spot, his failure to recognize that online commerce (access to "the long tail") is dominated by oligopolies.

Tim Wu is tougher in his *Slate* review (July 21, 2006): It's titled "The Wrong Tail" and subtitled "How to turn a powerful idea into a dubious theory of everything." As you'd expect from a *Wired* editor, even if you grant that Anderson's concept is original or applies to as many media as he claims, Anderson feels the need to carry it further. Wu says that when you

finish the book, "there's one question you won't be able to answer: When, exactly, *doesn't* the Long Tail matter?" As with so many business books, *The Long Tail* "commits the sin of overreaching"—including claims that offshoring is "the Long Tail of labor," online universities are "the Long Tail of education" and there's even a "Long Tail of national security." Most of which is just silly. Wu gets it right: The power-law curve (using the non*Wired* name) matters to business "1) where the price of carrying additional inventory approaches zero and 2) where consumers have strong and heterogeneous preferences." I'll suggest a third: Where it's plausible to *make* a large number of discernibly different products. 175,000 book titles in a year? No problem—especially with PoD. 175,000 different brands of dish detergent? Big problem: That's at least 174,000 more than could be differentiated. Wu mentions the oil industry: the power-law curve just doesn't apply. Even in information technology, the need for standardization frequently counts for more than heterogeneity. You don't want a choice of 10,000 different routers or *ten* different incompatible USB connection options; you want one that will work.

Nicholas Carr cites Lee Gomes at the *Wall Street Journal* in an indirect argument with Chris Anderson. Gomes claims the Long Tail's effects have been overstated. There was a critical change between the original *Wired* article and the book. The article says, "More than half of Amazon's book sales come from outside its top 130,000 titles." The book says, "About a quarter of Amazon's book sales come from *outside* its top 130,000 titles." When asked directly, Anderson admits that there are *no* current examples of sales of "misses" exceeding sales of "hits"—according to Gomes, Anderson says that won't happen for at least a decade at either Netflix or Amazon. So what? Well, Carr notes that Anderson's comparison (130,000 is roughly the stock of a typical Barnes & Noble) is flawed: "There have always been specialized bookstores, selling everything from religious and spiritual books to textbooks to foreign-language books to used and out-of-print books to poetry books... And there have always been small presses...selling books directly, through the mail." Unless we know how the power law functioned *before* Amazon, claims that the internet changed everything (in this case) are suspect. It's a new channel but not necessarily a change in buying patterns.

Does the power-law curve function in most media? Sure it does. That's neither new nor particularly

surprising. What's *somewhat* new is that the curve can keep trailing off to the right—the “long tail”—in TV (with more cable channels and short video on the web) and, more effectively, in movies thanks to Net-Flix. For magazines, there's nothing new here, although even smaller niches can be served entirely online. (According to some experts, most future growth in magazines will be in niche titles, but those have always represented more than 99% of magazine titles—there just aren't 2,500 mass-market magazines.) For books, it's not clear whether the internet makes the “long tail” more important. It is clear that most books have been niche books ever since thousands and tens of thousands of books came out each year. As for sound recordings, we have Big Media to thank for distorting the music scene in the 1960s through 1990s, dropping most artists to make way for bigger and bigger promotion budgets for a few Big Acts. Local clubs always acted against this “fat head” tendency (you could leave out the space there) and the internet acts further to restore a normal state, where thousands and tens of thousands of musicians work at smaller scale. Meanwhile, to be sure, Chris Anderson has the kind of Bestseller that supposedly doesn't exist in a “long tail economy.” I'm sure he's taking that irony to the bank.

Quicker Takes

The headnote for this edition talks about the handful of freely available unsponsored eJournals in the library field that existed in 2001, and the even smaller handful that continues in 2007. I didn't discuss *sponsored* free eJournals but those also come and go. I'm sad to say that one of the best, *D-Lib Magazine*, seems to be in jeopardy. *D-Lib* has been around for more than a decade. The January/February 2007 issue where “Current and future status of D-Lib Magazine” appears is Volume 13, Number 1/2. In the past, *D-Lib* has operated through government grants from DARPA and NSF, and more recently contributions by CNRI. That appears not to be a sustainable model. The magazine's moved to a bimonthly schedule for the next few issues, “at which point we will critically evaluate the options available to us.” *D-Lib* is trying to raise \$100,000 and may consider advertising or author charges. I'd hate to see *D-Lib* disappear. It's true that no publication can sustain itself without funding, whether “e” or print, particularly when editing is involved—and it's pretty clear that there aren't too many crazy people like Walt Crawford around (and I

wouldn't be willing to take on something like *D-Lib* as a *pro bono* effort: there's too much involved).

- It's always a breath of fresh air to encounter a cautionary note in *Technology Review*. One such appeared December 1, 2006: “What comes after Web 2.0?” by Wade Roush. The first paragraph's a tipoff: “Many researchers and entrepreneurs are working on Internet-based knowledge-organizing technologies that stretch traditional definitions of the Web. Lately, some have been calling the technologies ‘Web 3.0.’ But really, they're closer to ‘Web 2.1.’” Roush discusses the Semantic Web and notes the “gargantuan effort that would be required to tag all the Web's data with metadata” as well as narrower projects such as FOAF (www.foaf-project.org), Piggy Bank (simile.mit.edu/wiki/Piggy_Bank) and the Amazon Mechanical Turk, “a kind of high-tech temp agency” that pays low rates to do simple tasks computers can't handle. He concludes: “Most of these projects are so far from producing practical tools—let alone services that could be commercialized—that it's premature to say they represent a ‘third generation’ of Web technology.”
- A *long* time ago (December 7, 2005—the printout was mislaid), *Wired News* had a piece by Dan Goodin that probably still applies: “Old rips: May they rest in peace.” It starts with an undergrad who recently converted his CD collection (5,000 songs worth) to music files—for the *third* time. Why? “After spending years painstakingly compiling the perfect music library, he came to realize that the sound quality of the computer files left plenty to be desired.” He started with 128K MP3 then moved to 128K AAC, the format Apple uses for iTunes. But when he listened to CDs created from those files, he realized that something was wrong—even the more-efficient AAC loses too much sound quality at 128K. The article touts “Lame,” a particular MP3 encoding algorithm. I don't know much about Lame; I do wonder whether the primary benefit isn't using a higher bitrate. I've now ripped almost everything at 320K using MusicMatch's Fraunhofer MP3 compression...and that was after initially using 196K. I *know* I can tell the difference between 128K

and 196K (and between 128K and the original CD). I *think* I can sometimes tell 196K from 320K. I'm *certain* 320K always sounds at least as good (to me) as the original CD. ("At least as good"? Yes. There is a theoretical basis for believing that a copy could sound better than the original—in addition to euphonic distortion—but that's another topic covered previously, and I'm still not sure I believe it.)

- Paul Boutin wrote "Where's my Google PC?" at *Slate* on July 3, 2006. He believes it's coming—either a supercheap PC that's essentially a smart terminal to a Google-based set of services, or just a complete set of services with storage to match. But consider the tease: "It's coming. It'll be great. You'll hate it." That's oxymoronic, of course: It can't be "great" for you if you hate it. Boutin thinks it's "the future of computing" at one level—well, heck, he spent more than a decade "working on and around network-based computers and thin clients," so he's biased toward their advantages. He even claims network computing is *faster*, one of those "it depends" claims. He's pretty sure Google plans to build the "world's best network computer" and he's definitely hot for the idea. But he also knows why they won't work, even if he doesn't understand. "First, there's the inexplicable human urge to own stuff and have it in your possession." What makes that inexplicable? People *do* like to control their data. Boutin may think that's silly, but that's his problem. Second, a network computer only makes sense if you have a "fast, flawless network connection." Then there's the killer: "Are you going to let someone else handle all your data?" He will, "but that's my blind faith in statistics." Well, that and a long career-based bias toward network computing. His closing sentence says more than enough for me: "How about it: Would you trust Google to protect your e-mail, your tax documents, and your family photos?" Scratch that future, at least in my opinion.
- An interesting set of findings from Statistics Canada about internet users, cited on *slaw.ca* on August 2, 2006. Heavy internet users are avid consumers of other media, spending about the same time watching TV as nonusers—and spending *more* time reading books.

But heavy internet users "spent substantially less time in social contact with others." Maybe the term should be "browsing alone" rather than "bowling alone." Oddly, they seem to *enjoy* social events, clubs and organizations more than non-users—but they don't socialize much. Heavy users spend "considerably less time on paid work and domestic activities" and "less time to sleeping, relaxing, resting and thinking" than nonusers or light users. I leave you to add your own comments—and I'm guessing these findings would apply reasonably well in Canada's smaller (but far more populous) neighbor to the south as well.

- I'm not going to spend a lot of time on the podcasting study from Pew Internet, but it is worth noting a key finding and the significance of error margins. To wit: In a February-April 2006 study, 7% of internet users said they'd downloaded at least one podcast; a smaller August 2006 study showed 12%. One percent of users report downloading podcasts on a typical day. Now consider the margins of error: Plus or minus 2% for February-April, 3.5% for August. Did podcast usage increase between February-April and August? Not in a statistically significant manner, as far as I can see: the probable actual percentages overlap at 8.5% to 9%. If you said "around a tenth of internet users apparently use podcasts, but only about a tenth of that tenth do so frequently," I'd buy that. Any claims of rapid growth require more study.
- I ran into a lovely, lovely piece by Paul Di Filippo on his experience writing a *commissioned* article for *Wired Magazine*: "The joy of corporate journalism, by J. Ives Turnkey." You'll find it at www.pauldifilippo.com/glass_preface.pdf; the source page, www.pauldifilippo.com/articles.php, also shows the article Di Filippo turned in. You'll have to go to *Wired* to read the article as published. I'm a biased source here: I know Di Filippo from his sharply-written items in the *Magazine of Fantasy and Science Fiction* and elsewhere, so I'm predisposed to believe he turned in a good piece (reading it does nothing to dispel that belief). He tells a complex story about how writing gets turned into "supreme *Wiredness*,"

the homogeneous gee-whiz style that makes me happy my freebie subscription's about to expire. He cites six general trends in the many forced edits: All references to "the little people" were eliminated; ambiguity was minimized; facts were cloaked in "hipness"; the past was dismissed as unimportant; quotidian matters were de-emphasized ("boredom does not exist in the *Wired* cosmos"); drama was injected into basically undramatic situations. Sure sounds like the *Wired* I know and despair of, the magazine in which the future's always clear and techno-thrilling, the past doesn't matter and the only people are People Who Matter.

Net Media Perspective

Wikipedia Revisited

In previous episodes...

OK, that opening phrase appeared out of nowhere, but this may be a good time to revisit past items regarding the project.

2002-2004

The first mention came in May 2002 (*C&I* 2:7, p. 18), in a rare self-referential THE GOOD STUFF item citing a "PC Monitor" column in *Online* based on a survey done via *C&I*. (Circular enough for you?) After noting the article, I said:

If you read the print magazine (highly recommended) instead of an online version, look back two pages; Péter Jacsó's "pan" of the month is one of those grotesque "let's all make an encyclopedia" efforts (Wikipedia) that help some of us appreciate professional efforts. He is astonished that MIT's *Technology Review* ran a serious interview with the CEO and that Peter Suber wrote a friendly notice. I am incapable of being surprised by *Technology Review* behaving like *Wired*, but the Suber notice does surprise me.

Jacsó was indeed unhappy with *Wikipedia*—which was probably less than a year old when he wrote the column and had 16,000 articles at the time. Jacsó had "been panning in this column and elsewhere projects that I usually refer to as another encyclopedia on your lunch break," largely because the fancy titles "make even educated librarians and other decent people provide a link to these pathetic sites... Others will keep copying their links and thus adding clout to such sites ad nauseam—and that is not all right." Regarding *Wikipedia* itself, Jacsó calls it "a joke at best" and

notes, "It looks like a prank." Of the 16,000 articles, he notes that many of them were a single sentence and that more than 10% referred to September 11, 2001, "playing the emotional card, which is totally out of place in a general encyclopedia, and include information about lists of victims and vigil sites, personal reports—and scams."

Jacsó also mentions 25 articles about "the tiny island of Niue" and that most of them, like many other country entries, were lifted verbatim from a year-old edition of the *CIA World Factbook* (splitting each chapter of a profile into a separate article)—without crediting the source. He again called it "a prank" and included illustrations of portions of its "tips on contributing" and FAQ, showing "a lack of scholarly discipline" and a [casual] attitude toward quality information. He cites Peter Suber's surprisingly favorable comment (Suber called *Wikipedia* "the ultimate development in dynamic, interactive, collaborative scholarship") and concludes:

My, oh my, is this scholarship and ultimate? What would the naive users say and think, who will soon become contributors after reading the tips for contributors? Jimbo expects advertisers by mid-2002, and then you know who is going to be laughing all the way to the bank.

I cite this not to embarrass Peter Suber or Péter Jacsó. Peter Suber's comment included this key qualifier after "collaborative scholarship": "if you can call anything scholarship that dispenses with editorial filters in the name of user freedom."

Times change. *Wikipedia* has changed, leaving that 100,000-article goal in the dust. Some articles are impressively long. Some appear well based in cited sources. For many contemporary topics, *Wikipedia* is a wonderful starting point. The sin even back then in quoting big chunks of the *CIA World Factbook* was not copyright infringement (it's government-prepared, public domain in the U.S.) but plagiarism: failure to cite the source. Such plagiarism continues, although many articles are heavy with footnotes. "Jimbo" decided *not* to run advertising on *Wikipedia* itself. The site no longer looks like a prank.

I'm nervous about calling *Wikipedia* "collaborative scholarship" or scholarship of any sort—particularly given *Wikipedia's* rules, which forbid inclusion of original scholarship. It certainly has clout; *Wikipedia* articles appear at or near the top in many search engine results. (Second for "Niue" at Google, *first* numbered result on Yahoo! and Live—and, remarkably, tenth at Ask.) In 2007 it's impossible to dismiss

Wikipedia as a prank or joke—but it’s still controversial. Even *Wikipedia*’s entry on “*Wikipedia*” notes controversy, including this paragraph “above the fold” (before the outline of a fairly long article):

There has been controversy over *Wikipedia*’s reliability and accuracy, with the site receiving criticism for its susceptibility to vandalism, uneven quality and consistency, systemic bias, and preference for consensus or popularity over credentials. Information is sometimes unconfirmed and questionable, lacking the reliable sources that, in the eyes of most regular contributors, are necessary for an article to be considered of high quality.

I didn’t mention *Wikipedia* at all during 2003. Then came October 2004 (*C&I* 4:12, pp. 2-4) and the relatively brief PERSPECTIVE: WIKIPEDIA AND WORTH.

Late summer saw a whole bunch of foofaraw about wikis and specifically *Wikipedia*. After one columnist suggested *Wikipedia* as a resource for computer history, other writers assaulted *Wikipedia* as worthless trash; at least one librarian made noises about the difference between online junk and authoritative sources; some wiki advocates pontificated about the awesome error correcting capabilities of community-based collaborative media. Alex Halavais of the School of Informatics at Buffalo University made 13 changes in the English language *Wikipedia*, “anticipating that most would remain intact and he’d have to remove them in two weeks.” Presumably, if that had happened, there would have been evidence that the ease of modifying *Wikipedia* makes it suspect as a resource.

I discussed the results of the Halavais test (which impressed Halavais himself) and two commentaries (one of which concerned wikis and scholarship but not *Wikipedia* itself). One commentary was by Ed Felten, who reviewed five entries on “things I know very well” and found four of them good—but the fifth “riddled with errors” of a sort that would “lead high-school report writers astray.” Felten mentioned high school; as with most other college faculty, he would (I believe) assume college students would *never* use *Wikipedia* or any other encyclopedia as a cited source. Given one of the issues raised later in this PERSPECTIVE, it’s worth noting that the error-filled article related to Microsoft. I offered my own middle-of-the-road perspective, doubting *Wikipedia* would “eclipse” traditional encyclopedias and assuming it was neither worthless nor better than a traditional encyclopedia—and that entries should be used on a “trust but verify” basis. In the next issue (November 2004, *C&I* 4:13, pp. 14-15) I cited some direct feedback and related list discussion, including Michael Lorenzen’s cogent comments about *Wikipedia*’s extensive use of

government-generated public domain information creating a pro-American bias. My take then was that triumphalism was a problem with *Wikipedia* and its advocates—the felt need to “sweep away” or be “better than” traditional encyclopedias. That’s still true.

2005-2006

I mentioned *Wikipedia* in five 2005 issues of *Cites & Insights*, but some of the mentions were trivial. February 2005 (*C&I* 5:3, pp. 11-19) featured a new and considerably longer Perspective: WIKIPEDIA AND WORTH [REVISITED]. In case you’ve forgotten the situation in late 2004 and early 2005, three substantial critical essays appeared:

- 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 disagreed with the claims for collaborative editing and the methodology. He used one article as a case study, asserting numerous typographic, styling, grammatical and diction errors and calling it a “C [high school] paper at best”—despite more than 150 edits. He concluded: “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.”
- A few days later, Jason Scott posted “The great failure of *Wikipedia*” to one of his weblogs, *ASCII by Jason Scott* (ascii.textfiles.com). This computer history researcher “tried extended interaction with *Wikipedia*” and “consider[s] it a failure.” He says why in considerable detail. He argues with the project’s low barriers to entry, says it has “a small set of content generators, a massive amount of wonks and twiddlers, and then a heaping amount of procedural whackjobs” and comes to this judgment: “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.”

- On New Year’s Eve 2004, 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 /. The five-page article generated hundreds of comments very rapidly, the total back-and-forth reaching book length within the first month. Sanger was a cofounder of *Wikipedia* and continues to admire the project, which put him in an unusual position to point out what, in Sanger’s mind, has gone wrong.

I’m not going to recount the responses and list comments; the PERSPECTIVE is available at citesandin-sights.info/v5i3d.htm. I was surprised by Clay Shirky’s attack on librarians, teachers and academics (I said he “lost it”)—but since then, as Shirky has dismissed taxonomy as worthless and said web video will kill HDTV, I’ve become less surprised: Shirky is a man of strong, even strident opinions. Web4lib posts included extreme positions as well as nuanced discussion, as did Publib posts. My conclusion at that point:

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.

I discussed other commentaries in June 2005 (*C&I* 5:8, pp. 8-9), including a *Wired* essay with exactly the bias you’d expect, a disagreement between *Many2many* contributors and the continuing apparent need by some to create a zero-sum game, where *Wikipedia* can only “win” at the expense of traditional encyclopedias. A multitopic NET MEDIA roundup in October 2005 (*C&I* 5:11, pp. 6-7 for *Wikipedia* stuff) included notes on a long Larry Sanger memoir about the early history of *Wikipedia* and *Nupedia*.

While I mentioned *Wikipedia* four times in 2006, there was only one substantial discussion—NET MEDIA PERSPECTIVE: WHAT ABOUT WIKIPEDIA? (*C&I* 6:13, November 2006, pp. 2-11). By then, I was using *Wikipedia* as a starting point in many cases (it’s one of several choices in my Firefox search-box menu, along with Worldcat.org, IMDB and the four major search engines)—and I noted a case where a *Wikipedia* article directed me to a verifiable answer for a Unix problem

(where a book had provided bad information). “I don’t trust *Wikipedia*’s ‘neutral’ point of view and find many of the essays poorly written—but it’s great for what it is.”

The first major part of that discussion was *Nature*’s article comparing the accuracy of *Wikipedia* and *Britannica*. (I now see I managed to substitute *Science* for *Nature* once in that discussion, but I *certainly* don’t claim to be an authoritative source.) The study was, at best, too narrow to be conclusive—and at worst flawed both in its analysis and in the way it was reported. Perhaps the best commentary among those I discussed was Paula Berinstein’s “*Wikipedia* and *Britannica*: The kid’s all right (and so’s the old man)” in the March 2006 *Searcher*.

Then there’s notability and *Wikipedia*’s “inability to handle domain experts,” as danah boyd puts it. That discussion made one of *Wikipedia*’s true peculiarities clear (a peculiarity that’s mentioned in contemporary discussion, below): Living persons are *not* expected to edit their own entries except *perhaps* to correct factual errors. It’s “culturally inappropriate”—and although founder Jimmy Wales *did* edit his own article, he apologized for doing so. Seth Finkelstein discussed the impossibility of opting out: The astonishing case that, unlike (for example) *Who’s Who in America*, he can’t say “I don’t want to be in *Wikipedia*” even though he’s not a politician or otherwise so notable that he has no grounds for such a request. Others have noted the same problem.

I also discussed a lengthy *New Yorker* article on *Wikipedia*, a thoughtful piece that pointed out some of its strengths and weaknesses and included the pointed comment that “*Wikipedia*’s bureaucracy doesn’t necessarily favor truth.” A far less balanced piece appeared in the September 2006 *Atlantic Monthly*; it was a gushing, unbalanced tribute that *approvingly* says truth is whatever the community says it is. “Yes, that means that if the community changes its mind and decides that two plus two equals five, then two plus two does equal five.” As I noted, this model of “truth” would mean evolution is a myth, at least in America, since the majority of Americans (who respond to polls) apparently don’t believe in it. “And there’s no global warming and we’ll never run out of oil.” As long as we clap our hands long enough and loud enough, Tinker Bell can fly and our SUVs can keep getting bigger forever. A third mainstream press piece in the *Wall Street Journal* was an email debate between Jimmy Wales and Dale Hoiberg, current editor-in-chief of the *Br-*

tannica. It's an odd piece, considerably flawed by Wales' absolutism and dismissal of expertise.

Then there's *Citizendium*—Larry Sanger's fledgling project to produce a more authoritative *Wikipedia*. I discussed the early plans in that November 2006 essay, along with attacks from Clay Shirky and commentary from Nicholas Carr (who wrote off *Citizendium* before it got started). That project has changed course. More on that later.

We Interrupt this Program

Before discussing current controversies and developments in *Wikipedia* and *Citizendium*, it's worth noting again that very few wikis aim to be encyclopedic. A wiki is a tool—another kind of “free as speech” content management system, a bit less lightweight than a blog but better for different purposes. Sure, some people get overenthusiastic about wikis, creating them when simpler tools might work better. Certainly, there's a problem within librarianship in that people seem more inclined to create *new* wikis than to contribute to existing ones.

But *they do work*, sometimes exceptionally well. This PERSPECTIVE isn't about wikis in general, but I was taken with “wikilove,” a commentary posted January 24, 2007 by Jenica at a blog that's either called *Thinking out loud* or *Mermaid* (jenica26.squarespace.com/mermaid/). Jenica “love[s] our (password protected and unshareable) library wiki.” Excerpts:

One of my continuing frustrations is how to keep organized. As an area coordinator and a team leader, I have to organize information for my own use, I have to organize information so that our director can access it at the moment of need, I have to organize information for communal access by a committee, I have to organize information for communal access by a leadership team, I have to organize information so that the whole staff can find it, I have to organize information for the use of two working groups I chair. Each one of those information needs is different, and each one produces different kinds of information to organize. In some cases, the clear and appropriate solution is to use our shared storage drive. In some cases, we need to use a paper distribution system. In some cases, emailing documents is best. But I really hate those solutions for reasons of findability, retrievability, and convenience.

Which is why I love the wiki. Click on edit; edit; click on save; done. Click click click, retrieved. Or, search, click, retrieved. It's fantastic.

Sadly, it doesn't work for everything or for everyone, but as a finding tool it's fantastic...[Examples for some of

the cases above.] And since we use MediaWiki, there's a “watch this page” feature that sends me an email when my colleagues edit the pages I'm monitoring, which means that tracking the progress of a joint project just got effortless—the software does the checking for me...

It's a tool. Sometimes it's a remarkably effective tool. Jenica's instance isn't what some would call social software because it's protected—it's a shared resource for the staff.

Some wikis fail after initial excitement—just like most blogs. Some wikis generate new wikis more than they succeed on their own terms—just like some blogs. And some wikis become international phenomena and sources of ongoing controversy. But those are fringe cases that have little to do with the underlying uses and benefits of wikis.

In other words, criticisms levied at *Wikipedia* and *Citizendium* in the remainder of this essay are criticisms of *those particular wikis*, not of wikis as tools.

Recent Wikipedia Controversies

This one isn't recent except to me. “Digital Maoism: The hazards of the new online collectivism” by Jaron Lanier appeared at *Edge* (www.edge.org) on May 30, 2006. The blurb above the essay sets the tone:

The hive mind is for the most part stupid and boring. Why pay attention to it?

The problem is in the way the Wikipedia has come to be regarded and used; how it's been elevated to such importance so quickly. And that is part of the larger pattern of the appeal of a new online collectivism that is nothing less than a resurgence of the idea that the collective is all-wise, that it is desirable to have influence concentrated in a bottleneck that can channel the collective with the most verity and force. This is different from representative democracy, or meritocracy. This idea has had dreadful consequences when thrust upon us from the extreme Right or the extreme Left in various historical periods. The fact that it's now being re-introduced today by prominent technologists and futurists, people who in many cases I know and like, doesn't make it any less dangerous.

I'm not fond of the “hive mind” concept. The “wisdom of the crowd” strikes me as wildly overrated given, oh, elections, many online discussions, the general level of IMDB user reviews and other counter-examples. So what does “digital visionary” Jaron Lanier say about all this?

He starts out by complaining that Wikipedia identifies him as a film director. He's tried to fix that, but it doesn't stick. Now reporters are asking him about his filmmaking career (which consisted of one

experimental short film years ago). I won't summarize the entire essay. I will note a few items. Lanier says, "Accuracy in a text is not enough. A desirable text is more than a collection of accurate references. It is also an expression of personality." I couldn't agree more. To me, this is the difference between a bunch of facts strung together and a *story*. You could argue that collective editing works against story telling because it diminishes individual voices.

[M]ost of the technical or scientific information that is in the Wikipedia was already on the Web before the Wikipedia was started... In some cases I have noticed specific texts get cloned from original sites at universities or labs onto wiki pages. And when that happens, each text loses part of its value. Since search engines are now more likely to point you to the wikified versions, the Web has lost some of its flavor in casual use.

When you see the context in which something was written and you know who the author was beyond just a name, you learn so much more than when you find the same text placed in the anonymous, faux-authoritative, anti-contextual brew of the Wikipedia. The question isn't just one of authentication and accountability, though those are important, but something more subtle. A voice should be sensed as a whole... Even *Britannica* has an editorial voice, which some people have criticized as being vaguely too "Dead White Men."

If an ironic Web site devoted to destroying cinema claimed that I was a filmmaker, it would suddenly make sense. That would be an authentic piece of text. But placed out of context in the Wikipedia, it becomes drivel.

Lanier discusses the problems with "Meta" sites, sites that base their content on collective algorithms, in essence trying to reflect the "hive mind." Digg and Reddit are such sites, as is deli.cio.us to some extent, with popurls.com aggregating other metasites. He notes that popurls tends to ignore major but serious news (earthquakes, new approaches to diabetes management) in favor of pop culture and other trivia.

There are notions here I'd agree with, but there are problems as well. Consider this extract:

[I]t must at least be pointed out that writing professionally and well takes time and that most authors need to be paid to take that time. In this regard, blogging is not writing. For example, it's easy to be loved as a blogger. All you have to do is play to the crowd. Or you can flame the crowd to get attention. Nothing is wrong with either of those activities. What I think of as real writing, however, writing meant to last, is something else. It involves articulating a perspective that is not just reactive to yesterday's moves in a conversation.

That's not quite Gorman-esque, but still devalues many first-rate writers who blog (and who prepare their

posts seriously) and many skilled writers who do some or all of their writing without direct pay. "Blogging is not writing" is not thinking; it's unworthy of a professional writer or thinker.

Are there authentic examples of collective intelligence? Of course, and Lanier discusses them. He also asserts, "Every authentic example of collective intelligence that I am aware of also shows how that collective was guided or inspired by well-meaning individuals." In the end, Lanier's not entirely opposed to collective intelligence:

The hive mind should be thought of as a tool. Empowering the collective does not empower individuals—just the reverse is true. There can be useful feedback loops set up between individuals and the hive mind, but the hive mind is too chaotic to be fed back into itself.

Edge engages in "the reality club," where a bunch of Important Minds comment on something—in this case, Lanier's essay. The essay runs 12 pages; the responses take up 26 pages. I should note something about Lanier's essay that I probably wouldn't notice if I was a regular *Edge* reader (I tried it and decided against it), or if I was part of the In Crowd involved with *Edge*. To wit, Jaron Lanier makes sure that we know *he's important*. He's a big shot. He *matters*. Reporters talk to him frequently. The "prominent technologists and futurists" he disagrees with are "people who in many cases I know and like"—Lanier *moves in important circles*. Kevin Kelly "is a friend." Consider this sentence (after we learn that Lanier's a well-paid consultant who finds he's being paid just as much to do less these days): "I've participated in a number of elite, well-paid wikis and Meta-surveys lately and have had a chance to observe the results." I find this distracting and offputting. Clearly Lanier's prominent enough to be in *Wikipedia*. Does he really need to remind us so frequently that he's a bigshot and friends with People Who Count? If the essay is really about Jaron Lanier, maybe so. If it's about "digital Maoism," probably not.

I'll admit to ignorance here. Lanier is described as a "computer scientist, composer, visual artist, and author" (that description could be taken directly from his website). Worldcat.org shows zero books by Lanier, but that's OK: Lots of authors never get around to booklength projects. He claims to have coined the term "virtual reality"—and *Wikipedia* currently calls him a virtual reality developer. I don't doubt that Lanier is an important public intellectual. I *do* doubt the wisdom of his constant stressing of his importance

and connectedness. Incidentally, if you read the *Wikipedia* article on Jaron Lanier, make sure to read the Discussion page: To my mind, it says a little too much about what's wrong with *Wikipedia* and the way Wikipedians deal with true topic experts, including the legitimate expertise that living people have about their own lives.

I'm going to cop out. If only for reasons of space, I'm not going to attempt coherent comments on the responses from such luminaries as John Brockman, Clay Shirky, Douglas Rushkoff, Cory Doctorow, Kevin Kelly, Esther Dyson, Larry Sanger, Jimmy Wales, Dan Gillmor and Howard Rheingold (among others). Brockman informs us that we're migrating "from individual mind to collective intelligence" and "witnessing the emergence of a new kind of person"—and tells us how important all the other commenters are, including Clay Shirky, than whom "no one is deeper, more thoughtful, on the social and economic effects of Internet technologies." The commenters are "a 'who's who' of the movers, shakers, and pundits of this new universe of collective intelligence." Wow. Some of those movers and shakers write well; some, surprisingly badly. Some think well; some do not, at least in these commentaries. Some agree with Lanier (at least partially); some mock him and celebrate the "hive mind." After reading through all of the commentaries, I wrote a rude comment on the last page; I won't repeat it in this family-friendly publication. You may find the commentaries enlightening; I did not.

Plagiarism?

A November 6, 2006 Associated Press item by Anick Jesdanun notes a project by Daniel Brandt to check portions of 12,000 biographical articles in *Wikipedia* against other sources, looking for plagiarism—that is, uncredited copying. He brought 142 articles—just under 1.2%, for what that's worth—to *Wikipedia's* attention. Jimmy Wales called the findings "exaggerated" while admitting that plagiarism does happen. Frankly, if just over 1% of the articles in *Wikipedia* contain uncredited copying from other sources, I'd say it's doing pretty well.

Brandt is another person unhappy with his biography in *Wikipedia*. He runs Wikipedia Watch, Google Watch and Yahoo! Watch. I suggest looking at any or all of these sites; that certainly helped me form an opinion as to how seriously I should take Brandt's criticisms. *Ad hominem* may be a logical fallacy but it's useful in real life.

Can Wikipedia ever make the grade?

That's the title of an October 27, 2006 *Chronicle of higher education* article by Brock Read. Read covers some of the same ground I've noted and does so engagingly and skillfully, adding more sources and seemingly favoring neither *Wikipedia* nor its critics. I appreciate the comment of Roy Rosenzweig, a history professor at George Mason University: "Are Wikipedians good historians? As in the old tale of the blind men and the elephant, your assessment of *Wikipedia* as history depends a great deal on what part you touch." He finds "thorough, fairly well-written essays" on topics such as Red Faber and "Postage stamps and postal history of the United States," but "incomplete, almost capricious coverage" on possibly more important topics such as American history from 1918 to 1945 (which at the time omitted items such as female suffrage, the Ku Klux Klan and the rise of radio).

Read identifies at least three reasons scholars may be reluctant to add articles in areas where *Wikipedia* is currently weak: The difficulty of *keeping* scholarly work there when avid editors are quick to delete or dumb it down; the fact that anonymous articles in *Wikipedia* won't do a thing for scholarly respect or career advancement (why write for *Wikipedia* when you can be working on refereed articles?); and *Wikipedia's* preference for concise articles (which seems less in evidence for pop culture). He has three scholars grade individual *Wikipedia* entries, with results ranging from A (for "Flow cytometry") through C (for "African-American civil rights movement").

The article's accompanied by a sidebar and, on *CHE's* website, a "live discussion" involving Alex Halavais and a number of questioners. Halavais is good: He believes college students shouldn't be allowed by their professors to cite *Wikipedia* in research papers—"but only because it is an encyclopedia" (a stance with which founder Jimmy Wales agrees). Halavais now seems to be a *Wikipedia* supporter, based on his responses in general. Marc Meola commented on the discussion in an October 27, 2006 post at *ACRLog*. He didn't change his basic view, "which is that the errors are too random and the editing too chaotic."

Wikipedia and the trust factor

Paul Valley wrote this piece in *The Independent* (London) on October 22, 2006; I picked it up from *Tech-NewsWorld*. Valley asserts bias in a number of *Wikipedia* entries because the work of "dedicated contributors with idiosyncratic beliefs" sticks around—

“because no one has the time and energy to counteract them.” In some cases, “pages seem to have been taken over by fanatics and special interest groups.” Valley also notes “disproportionate emphasis,” one of the more cogent criticisms. Valley isn’t anti-*Wikipedia* but he is cautious. I like this interchange:

How unreliable is it?

How long is a string of clichés? That’s how a lot of *Wikipedia* entries read. But then others read as if they were written by people who know what they’re talking about. The problem is all the stuff in between, which looks reliable, but you never know. Using *Wikipedia* is like asking questions of a bloke you met in the pub. He might be a nuclear physicist. Or he might be a fruitcake.

Valley’s conclusion: “*Wikipedia*’s premise—that continuous improvement will lead to perfection—is completely unproven.”

Knowledge and unknowledge

Nicholas Carr posted this on December 3, 2006 at *Rough type*. Carr blogs about *Wikipedia* a lot. This time he notes that “elite members” at *Wikipedia* increasingly spend their time removing stuff, including real content. An article on a Canadian band was removed because the powers that be decided the band was “non-notable,” drawing a razzberry from the bandleader. According to Carr (citing a *Washington Post* article), about a hundred entries a day bite the dust—and Carr wonders why.

Now, philosophically, I have no problem with this new-found desire to separate the wheat from the chaff. Encyclopedias have always had to decide what’s worthy of being included and what isn’t. *Wikipedia* is just following the fine old tradition of selectivity. But what puzzles me is this: I thought *Wikipedia* was about *not* following tradition. I thought it was about being freed from the old physical world’s scarcity-imposing constraints, the constraints that forced us for millennia to live without easy access to “the sum of all human knowledge.” I thought the fact that *Wikipedia* didn’t have to worry about ink and paper and printing meant that it could be radically inclusive—that it could put everything in and let readers decide what was worthy of their time and what wasn’t. I thought *Wikipedia* was about the long tail of knowledge. I thought it was about *abundance*.

He sees a discrepancy between ideal and practice and notes a response by Mitch Kapor to a question about gatekeeping: “Who said that quality emerges out of gatekeeping?” Carr’s comment: “Who said that quality emerges out of gatekeeping? That’s precisely what *Wikipedia* is saying, about a hundred times a day.”

In comments, Seth Finkelstein thinks Carr has put *Wikipedia* in a no-win situation and notes there

have always been *some* standards for inclusion, trying to avoid “things that nobody will care about.” Michael Moneur believes *Wikipedia* has “been publishing volumes upon volumes of things nobody cares about for years,” such as pages about characters in obscure videogames and *fans* of characters in obscure videogames. Anthony Cowley agrees *Wikipedia* is violating its own slogan, “the sum of all human knowledge.” It’s an interesting and inconclusive discussion—but, much as I admire Seth Finkelstein’s work, I come down on the other side here. *Wikipedia* is absurdly inclusive on things geeks care about—pages on individual *Mutant Ninja Turtles*—and far more selective on “minor” topics that its core contributors and activists don’t care about.

Wikipedia will fail in four years

That’s the somewhat startling title of a December 5, 2006 post by Eric Goldman in his *Technology & marketing law blog*—and I should note that Goldman *does* have a (stub) entry in *Wikipedia*, with its own typically-interesting discussion page regarding his notability. This essay is actually a year-later update of a prediction that it would fail in five years. His key points:

- Growing traffic makes *Wikipedia* a target for marketers. Wikipedians are the only things keeping marketers from screwing up pages.
- Marketers will use automated tools to attack pages, forcing Wikipedians to spend more time and energy combating them, leading to burnout and causing them to leave the project, leaving fewer people to pick up the load.
- This will lead to increasingly junky pages—which may lead to more Wikipedians bailing out because it’s no longer worthwhile.
- “Thus, *Wikipedia* will enter a death spiral where the rate of junkiness will increase rapidly until the site becomes a wasteland. Alternatively, to prevent this death spiral, *Wikipedia* will change its core open-access architecture, increasing the database’s vitality by changing its mission somewhat.”

Goldman uses the Open Directory Project as an analogy, noting that ODP in its heyday did an “amazing job of aggregating free labor to produce a valuable database”—but ODP is “now effectively worthless.”

How’s it going a year later? There are still relatively few active editors. Jimmy Wales says the vast majority of work is done by about 1,000 Wikipedians.

Goldman suggests you could argue *Wikipedia* has already diverged from an open-access paradigm, becoming more insular and self-focused. There aren't a lot of rewards for being an active Wikipedian: it's free and largely anonymous labor. Goldman believes marketers are increasing their pressure on *Wikipedia*. He stands by his prediction.

The Microsoft controversy and Wikipedia's expertise problem

Microsoft was unhappy about *Wikipedia's* entry on the Open Office XML document format. Someone at Microsoft offered to pay an independent expert to edit the entry. The offer was not hidden. The expert was free to write anything he wanted. Here's how it was described in AP coverage on January 23, 2007, quoting Microsoft spokesperson Catherine Brooker:

Brooker said Microsoft had gotten nowhere in trying to flag the purported mistakes to Wikipedia's volunteer editors, so it sought an independent expert who could determine whether changes were necessary and enter them on Wikipedia. Brooker said Microsoft believed that having an independent source would be key in getting the changes to stick — that is, to not have them just overruled by other Wikipedia writers.

Brooker said Microsoft and the writer, Rick Jelliffe, had not determined a price and no money had changed hands—but they had agreed that the company would not be allowed to review his writing before submission.

Jimmy Wales' response? "Wales said the proper course would have been for Microsoft to write or commission a 'white paper' on the subject with its interpretation of the facts, post it to an outside Web site and then link to it in the Wikipedia articles' discussion forums."

To which I can only say, "Huh?" It's unacceptable for Microsoft to submit material on an area it should be expert on *and* it's unacceptable to hire an expert to do so—but it *would* be acceptable to post the material somewhere else in article form, then link to it. After all, *Wikipedia* material can't be "original"—but cited sources can be almost any web site, including blog posts and /. essays. Even then, the expert isn't allowed to make corrections, only to call for discussion.

I think Nicholas Carr's take on Wales' comment is worth quoting:

That's kind of an odd suggestion from "the encyclopedia that anyone can edit." It seems like we're getting to the point where anyone who has gained deep enough knowledge of a subject to have developed a point of view on it will be unwelcome to edit Wikipedia. Experts, automatically considered suspect, will be forced to go through some parody of a traditional editorial process.

Seth Finkelstein posted "Wikipedia articles can be a disaster waiting to happen" on January 24, 2007 at *Infthought*, noting the Microsoft controversy and the extent to which *Wikipedia* "can be a minefield of conflicting rules, administered by petty bureaucrats, with a collection of obscure policies that spawn the term 'wikilawyer.'" As Finkelstein notes, this isn't always about ego. In the Microsoft case, where data formats are being described, "there's big bucks at stake." I love the closing paragraph:

Maybe this specific argument just comes with the territory of money and power. But still, it's quite a feat to make me feel sympathy for Microsoft.

Later that month (January 29, 2007), Finkelstein noted posts he calls "Wikipedia punditry," most having to do with *Wikipedia's* mysterious internal bureaucracies, problems with expertise and internal power-tripping. Chris Edwards notes that the Byzantine procedure for suggesting a change when you're an expert is "surprisingly close to that used by traditional publishers... That is, it would be if the publisher had a bureaucratic system based on China's [local bureaucracies]...each one does it differently, and attitudes can change dramatically in the space of days, although they will refer to the same rule book and come back with some obscure answer..."

I was particularly interested in Kathryn Cramer's detailed proposal (January 25, 2007, www.kathryncramer.com/kathryn_cramer/): "SF author bios should be moved from Wikipedia to the ISFDB wiki." Currently, the Internet Science Fiction Data Base (an essential resource for SF anthologists like Cramer) does *not* include author bios; it relies on *Wikipedia* for such bios. Cramer thinks that's the wrong way around:

After a brief experience with Wikipedia, its editors strike me as a pack of officious trolls whose main concern is to make sure that you don't actually know the people you are writing about. The science fiction field doesn't work that way. I know hundreds (maybe over a thousand) science fiction writers, editors, and fans. Many, many of them could be described as my "associates." Am I connected to most members of the professional science fiction community in some way? You bet.

I've helped run a Hugo-nominated SF semiprozine for a couple of decades, I edit two year's best volumes, and am married to one of the most eminent editors in the field. But this connectedness holds true of really a lot of the people doing the actual biographies: Perhaps their connections are not so visible or so obvious, but the SF field is like one big extended family. We've all slept on each other's couches. We've bought each other drinks. We marry each other's daughters... It's Clan Fandom.

And of those creating biographies that don't know their subjects, what they are mostly doing is lifting the ISFDB bibliographies wholesale and transplanting the content over to Wikipedia.

Cramer's asserting, I believe correctly, that *Wikipedia's* distaste for being too close to your subject *just doesn't work* in science fiction and fantasy: The writers and fans have always intermingled, and as "semiprozines" show, the lines are fuzzy.

It's a long post and makes an excellent case. The best critics in science fiction tend to know the authors. That doesn't cause them to pull punches—but it does, apparently, make them ineligible to write *Wikipedia* bios. She cites examples of well-written bios deleted or rewritten to eliminate material based on personal knowledge—substituting the bland stuff any Googler can find.

The proposal's legal, as long as the ISFDB wiki has an appropriate license, but it's clearly just a starting point: ISFDB should have *real* bios written by knowledgeable people. In updates, Cramer notes the Microsoft controversy and Jimmy Wales' astonishing letter saying that editing *Wikipedia* for pay is unethical and "a grave violation of community trust." Cramer says the controversy "reflects more on problems with Wikipedia than with Microsoft; Wales' own attitudes promote the kind of bureaucratic paranoia and suspicion of expertise I experienced... Truth is not the point. The point is control." And this:

To me the biggest irony of the Microsoft controversy is that the material that the Wikipedians I talked to insisted was the only kind of material appropriate for sourcing was pretty much all written by people who were paid to write it. And edited by people who were paid to edit it.

Comments on the proposal begin with a charming irony: If ISFDB *did* create its own bios, "Wikipedia could then happily lift all the information from the ISFDB bios, and they would be properly 'sourced.'"

Citizendium

Citizendium hasn't disappeared; in fact, it's now opening registration. But there's been a big change in the project. Based on comments from early contributors, Larry Sanger concluded that forking *Wikipedia* in its entirety resulted in too much mediocre material, which contributors found tiresome to edit.

So the fork is being "unforked." All *Wikipedia* articles in *Citizendium* that haven't yet been edited by *Citizendium* contributors will be (or have been) deleted. That makes the site much smaller but should

substantially improve the average quality of entries—if Sanger's basic theses are correct. It's far too early to tell whether that's true.

Comments and Conclusions

When Jimmy Wales says college students shouldn't cite *Wikipedia* in research papers because they shouldn't cite *any* encyclopedia, I agree. When Jimmy Wales says, "One aspect of Jaron Lanier's criticism had to do with the passionate, unique, individual voice he prefers, rather than this sort of bland, royal-we voice of Wikipedia. To that I'd say 'yes, we plead guilty quite happily.' We're an encyclopedia," I disagree. Lanier struck me as calling for *voice*—not necessarily "passionate" but coherent, turning sets of facts into stories. There is *nothing* about an encyclopedia that precludes coherent, well-written entries representing single voices with personality; groupthink and bland speech are not prerequisites for encyclopedia entries. (Remember the "scholars' edition" of the *Britannica*?)

I looked up an article in a traditional encyclopedia (albeit one in DVD form): *Encarta 2007*. The article, "Pre-Columbian art and architecture," is long, segmented, and *interesting*; it's written in a clear voice that tells a story. It's also *signed*, in this case by Robert J. Loscher of the Art Institute of Chicago, an expert in the field. So are many articles in many encyclopedias. Wales' defense is simply nonsense.

One frequently cited issue, the uncertainty as to whether stuff in *Wikipedia* has any basis in fact, is to some extent being dealt with as articles show ever more footnotes. Unfortunately, that process seems to have two negative side effects: It makes the articles harder to read (when there are superscript numbers every sentence or two), and it may be making articles even less coherent and "voiced."

There's nothing earthshaking here, just an update on continuing issues in the wacky world of *Wikipedia*. I believe *Wikipedia's* "all human knowledge" claim has been sufficiently revealed as hypocritical to be discarded. One editor's "non-notable" is another person's substantial. Will *Wikipedia* die (of bloat, marketing or other problems) by 2010? How should I know?

I'm happy I *don't* have an article in the English *Wikipedia* (there is, as of this writing, a stub in the German version). I'll do more notable people in the library field the favor of not creating articles on them either. Increasingly, the "solutions" to *Wikipedia* problems appear to be problematic—and for living subjects, inclusion in the project can be the kind of honor

one might wish to duck. But while you can opt out of *Who's Who in America* or *Who's Who in the World*, you can't opt out of *Wikipedia*—unless, of course, you're part of the mostly-pseudonymous inner circle. And that's just wrong, unless you're a politician, Nobel laureate, before-the-title actor or similarly public person.

My Back Pages

All Media are Ad-Supported?

I don't want to pick on Steve Smith or his "follow the money" column in *EContent*. He frequently has good things to say. But in his September 2006 column, "Commercializing mobile content," he stumbles:

From newspapers and magazines to television and radio, history shows that advertising helps make a new platform cheap enough to achieve mass penetration and evolve a lucrative ecosystem. Video-on-demand, TiVo, and satellite radio are all good examples of cool technologies that have yet to break into large audiences, and all of them rely on fee-based models.

So *that's* why books, records, and movies never reached large audiences! They don't rely on ads as their revenue model. Why, if only books and records were free and had lots of embedded ads, people might actually read and listen. Live and learn.

Resistance is Futile

That's the title of Barry Willis' "The long view" column in the Winter 2006 *The Perfect Vision*. He's talking about chipping. RFID chips, that is. **In people**. He's all for it.

Why? You can have your "smart home" track you so that the lights are on and your music is playing, but only in the room you're currently using. Willis describes a "home of the future" where the tracking depends on a Bluetooth-enabled cell phone. "And that's the problem. Cell phones can get lost." Willis thinks RFID chips are better—and hey, vets are putting them in pets, aren't they? Here's the conclusion:

Much ado about nothing? Perhaps. For the 21st century technophile, nothing could be more convenient, or appropriate. The whole thrust of modern technology is toward deeper integration of humans with systems and content. A biotech interface is simply the next step. As *Star Trek's* Borg so ominously intone, "You will be assimilated." What they don't say is how much you'll like it.

Not for me, thanks. I'm really not planning to become more deeply "integrated" with "systems and content,"

for that matter. Maybe you're ready to be chipped up (not just at the poker table), but I'll pass.

Technoride?

Making fun of *PC Magazine's* three pages of auto coverage each issue is like shooting fish in a barrel, but even so... The hot item for November 7, 2006: A full-page blurb for the GMC Yukon XL as "an SUV that makes sense." There must be situations in which an 18-foot-long 2.8-ton behemoth (that doesn't carry passengers as well as a minivan) "makes sense," but I'm not sure how this beast qualifies as a technological powerhouse. Maybe the tech is "active fuel management" (shutting down half the cylinders at highway speeds) or the ability to use either E85 (85% ethanol) or regular fuel. But the mostly-highway test-driving yielded "mid- to high teens": In other words, not even *close* to the claimed 21mpg. As for E85, the "handwritten" caption in the photo touts switching from gas to E85 as "greatly improving fuel economy"—while the text, in two different places, says E85 *cuts* fuel economy by 25%: "You'd be doing well to see 15mpg." On the highway. Ain't tech wonderful?

A Eulogy for Old Media

I just love posts where the writer reaches conclusions then writes the story to support the conclusions *no matter what the facts show*. Scott Karp wrote "A eulogy for old media" at *Publishing 2.0* on August 21, 2006, and starts out stating a conclusion: "A eulogy is a speech of praise, typically—although *not necessarily*—for the dead, which seems fitting for a post about the lingering charms and strengths of Old Media."

You get "dead" and "lingering," right up front. Followed by survey results saying that people trust old media a lot more than they do the net—and that most people don't watch video online and have never paid to download video. Two-thirds never use social networking sites. 71% have never posted a comment on a blog. 79% have never written a blog. Only one percent of cell-phone users choose service providers based on entertainment options.

But Karp focuses on convergence, so he's got an answer: "We're still just seeing the tip of the iceberg." "So New Media is indeed a force to be reckoned with and will eventually take a huge share of attention from Old Media." Why? Because Scott Karp says so. (He goes on to admit that New Media *won't* kill Old Media—it doesn't work that way. In other words, he's writing a eulogy for something he admits isn't dying.)

Flicker: It's a Good Thing

I was bemused by a portion of a reply to a letter to *Home Theater*, appearing in the January 2007 issue. The correspondent wondered whether there could be TV sets that would scan at 24 frames per second (fps) rather than 30 (or 60) fps. It's a good question because movies (at least those made on film) are filmed at 24fps; that has to be translated to DVD or TV using a "3:2 pulldown" process, which can cause problems.

The answer was, in part, that some displays do indeed use multiples of 24fps, typically 72, 96, or 120; 24 or 48 would be too slow and would cause visible flicker. Then it gets interesting:

Keep in mind that, while eliminating the 3:2 sequence is great, it doesn't mean the image will be smooth. Film (24 fps) will never look as smooth as video (30 fps), as the difference in frame rate is noticeable.

This is a good thing. You don't want your movies to look like your evening news (unless that was the intention). A 2:2 conversion (or 3:3, or 5:5) will be smoother—as in less jerky—but not perfectly smooth.

Got that? Movies are *supposed to be* jerky: That's how we know they're not TV. Maybe you could have simulated light leaking around the edges, little blobs on one corner of the screen every X number of minutes, a faint whirring sound in the background—that would *really* simulate the big screen.

Really Long Mortgages

The January 2007 *Perfect Vision* includes a report on the 2006 CEDIA Expo—think custom home theater. One section discussed Exceptional Innovation of Columbus, Ohio, which puts together whole-house control systems. "The oven can interrupt your Web-surfing...to remind you that dinner is ready." Here's the key quote:

A complete EI system installed in a new home should run about \$80,000, conceivably rolled into a mortgage ("About \$65 per year," said one EI exec). Labor-intensive retrofits will cost more.

I'm not going to suggest that \$80,000 for a *control system* for a house is outrageous (more if you're adding it to an existing house); it's not my money. I am going to suggest that repaying \$80,000 at \$65 per year as part of a mortgage is, um, tricky. Let's see: At 6.5%, that \$80,000 will cost more than \$4,000 per year in added interest, and you're paying it off at \$65 a year...well, people with \$80,000 to spend on home automation can probably make better sense of those numbers. Change the figure by two orders of magni-

tude and it might work: At \$6,500 a year and 6.5% interest, the \$80,000 would be paid off in 25 years—and you would have paid another \$81,000 in interest. Quite a price for extreme interactivity.

I'm also charmed by the writer's "Best of" list, which suggests all of us will be hot for this sort of thing. To quote:

Greatest Technological Breakthrough and Most Significant Industry Trend. Deep whole-house integration as demonstrated by Exceptional Innovation. Extreme in degree of interactivity and multiplicity of options, it's technology for the multitasking, short-attention-span generation: fascinating, disturbing, and ultimately where the 21st Century is headed.

Good thing there will never be energy shortages or global warming and people all have high-paying jobs that suit their short-attention-span lifestyle.

Free and Easy—and Legit

My freebie subscription to *Business 2.0* is almost over. I'm delighted, particularly given the "web page" redesign of the magazine—but mostly given its consistent "the only ethical issues in business are whether you make money and whether you are provably criminal." Spam link pages? Hey, you can make a buck. Domain ranching? Wow, profit ahead. But that's a topic for a blog post one of these days.

This one's a "What's Cool/Playing the Angles" one-pager, "Free and easy," in the December 2006 issue (before the big redesign, when the magazine still had serif type and justified margins, and lacked hot buttons on every page). It discusses Mat LaClear's website, selling four books on real estate. The first one, *No Holds Barred: Mugging Tactics for Today's Real Estate Agent*, costs \$50 and made LaClear \$15,000 in its first month. "But here's the rub: He didn't even write it." It's a slightly revised version of *Closing the Sale* by J.C. Aspley, a 1925 title that apparently didn't have its copyright renewed. (Anything before 1923 is automatically fair game in the U.S.)

LaClear finds books in the public domain that might be salable today, possibly with a little reworking. He scans the book, does a bit of updating, adds a new title and puts it out as a new book under his name. He's not the only one reselling public domain material; the story mentions several others (and Dover used to publish a number of books entirely based on public domain material). The key to LaClear's success appears to be targeted email lists—that, and the ap-

parent willingness of would-be real estate hotshots to pay big bucks for these how-to books.

Here's the final paragraph:

Easy—but what about sleazy? The J.C. Aspleys of the world might bristle at content resellers basically plagiarizing their work for profit. The resellers don't see it that way, of course. "Is it fair to mankind that good, helpful manuals are lost forever?" LaClear asks. "I view myself as a recycler."

So do I—although LaClear really should include a back-of-title-page note that "Portions of this book originally appeared in different form as *Closing the Sale* by J.C. Aspley." Claiming authorship for the "new" book if he's just rewritten a couple of anecdotes is sleazy, and would be entirely unacceptable for scholarly work.

But reusing public domain material in new works is neither illegal nor unethical, and I wouldn't call it sleazy if appropriate credit is given. The public domain's *supposed* to generate new content and new uses for old content.

You Need All 107!

Another fish in a barrel: "Wired tools 2006," a 77-page feature in the December 2006 *Wired* (another freebie that's almost expired) featuring "107 cool upgrades to your holiday wish list." The first page says, "We're convinced you'll want to get your hands on *all 107 cool tools* in this guide. We do." [Emphasis added.]

How could there be a list of 107 "tools" that any one of us would desire, particularly the kind of "tools" *Wired* would feature? It's a puzzler. So I waded through the enormous advertorial—er, journalistic masterpiece heavily interlarded with real ads.

Start with a \$700 cell phone—but you *also* want two "\$TBD" cell phones (let's assume \$200 each), three more phones (\$349 total, but you're signing lots of two-year contracts at this point) and a \$200 BlackBerry—plus \$152 worth of accessories. Maybe the average *Wired* reader really does to own six new cell phones and a new BlackBerry, and after all they're only out around \$1,800 so far (plus usage contracts).

The next category is home theater, where you start with a \$19,995 (call it \$20K) front projector and add four TVs (one **\$70,000**, one a mere \$11,000, a third only \$5,000, and a cheapo at \$4,299), plus *another* front projector—but that one's only \$14,999. What *Wired* reader doesn't need four new big-screen TVs and two front projectors, all for a mere \$125,293? Note that "home theater" apparently

doesn't include surround sound: There's not an extra speaker in the roundup.

"Toys" I could see wanting and paying for: A \$150 microcontroller kit, \$20 "big bad booming bugs," \$160 spy video car, \$120 flying boat, \$999 robot, \$300 fur pony, \$70 Teddy Ruxpin, and four chemistry sets totaling \$90. Heck, given the apparent tastes and income of a proper *Wired* subscriber, maybe their deserves at least \$1,909 worth of toys. We're up to \$129,000 or so.

Here comes audio: A \$460 dynamic effect/sampler, \$12,700 "rotary woofer," \$1,999 Onkyo speakers (only two!), \$250 Zune, \$250 streaming controller, \$100 noisemaker, \$1,300 receiver and streamer, \$13 earmuffs, \$25 "speakers" (one watt! hey, hot stuff!). Mostly gadgets, suggesting that *Wired* folks care a lot more about TV than music. \$17,097; we're up to roughly \$146K.

Next come cameras. Of *course* every *Wired* subscriber's a good enough photographer to justify a Leica M8 digital SLR (\$5,300, not including lens)—but you apparently also want *six more* still cameras and a digicam that stores video on SD cards. That's seven still cameras, one for each cell phone; along with accessories (but not including the lens you'd need to actually take pictures with that Leica or another body-only camera), the camera package runs \$11,123. Chump change, taking us to around \$157K.

Sports time. \$1,125 skis, \$1,999 kayak, \$199 "digital coach," \$75 turtleneck, \$235 golf club (one club), \$30 "hydration monitor" (a high-tech water bottle!)—I have no argument here, since there's only one of each and the total's only \$3,663 (taking us a little over \$160,000).

Video: The \$450 Archos 604 portable player, \$250 Slingbox Pro, \$1,000 Linux-based DVR, \$8,000 one-terabyte media center, \$1,000 Toshiba HD-DVD player, and \$60 indoor ATSC (high-def) antenna. Not at all implausible, although the \$8,000 media center seems mildly overpriced. This group totals \$10,510; running total—call it \$170,000, rounding down. After all, at this point, what's a kilobuck or so?

Not enough to buy the first Household item, that's for sure: the Sheer Kitchen, a cooking station going for \$99,000. But wait! There's more! \$1,699 wall oven, \$599 dishwasher, \$6,795 *door* (it's a steel door), \$176 really ugly plastic chair, \$150 handheld vacuum. That's \$108,419 for household goods; we're up to more than \$278,000.

We had toys before. Now we've got games: \$250 Wii, \$50 guitar controller, \$130 gaming mouse and pedals, \$100 or more for the HyperScan Game System, \$280 "Killer Network Interface Card," \$30 card game, and around \$400 accessories. \$1,240 or so.

That's the lot. Can you count 107? Yes, with all accessories. You can have it all for roughly \$280,000. I'm thinking about the kind of household where putting a \$70,000 TV or \$99,000 prefab kitchen on a Christmas wishlist seems plausible. Maybe I'm just not in *Wired's* desired demographic.

Whaddya Expect for \$43,440? Quality Control?

High-end stereo pricing has more to do with rarity and perception than it does with value for money. Sometimes it gets a little crazy, though. Kevin Thornton noted that in a letter to *Stereophile* (in the January 2007 issue), referring to a very favorable review for a CD transport and digital/analog processor (basically, a multipiece CD player with some added capabilities). Although the reviewer loved the unit, the measurements were awful.

The manufacturer's response to the awful measurements? Oops, the unit wasn't wired properly. We'll send you a new one: It should measure *much* better. And while *Stereophile's* practice is to review the unit sent, they will do updates in cases like this—and they don't dwell on the fact that a manufacturer sent a defective unit to a reviewer.

These two units—basically a fancy multibox CD player—cost \$43,440. Fortythree big ones and change (plus another three kilobucks or so for sales tax in some states). Thornton finds this a little outrageous:

I'm sorry, but at the combination's price of \$43,440, one expects, and should get, perfection. How did this unit leave the factory with faulty wiring?... It amazes me when *Stereophile* reviews an insanely expensive product and it turns out to have a manufacturing defect. If manufacturers aren't fussy about what they send reviewers, what chance do we consumers have?

Good point. (Another good question: How come a golden-ear reviewer, who also believes a \$99,000 turntable is worth buying because it sounds so great, *couldn't hear* that the unit was seriously defective?)

Toward the end of the January issue there's a followup on the Halcro dm38 stereo power amplifier, where the reviewer found more distortion than expected. The manufacturer said "Oops. A part was out of tolerance" and sent a new one. Which measures

better. This time it was a small capacitor; shipping damage seems unlikely to be at fault. Now, the Halcro costs a mere \$18,790; maybe it's *especially* unreasonable to expect quality control at such a "low" price!

Making Print Look like the Web

I've commented before about books that look like HTML dumped to the page—with blank lines between paragraphs instead of first-line indentation, excessive leading, no distinction between paragraphs below headings and other paragraphs and so on. Well, *Business 2.0* took it a little further with a 2007 redesign. The editor's delighted:

Every page is shiny, colorful, and covered with widgets (key "take-away" information about the story), creating a magazine that looks more like a Web 2.0 site than a traditional print publication. That's deliberate. "The magazine's design has always been intentionally simple and clear," [designer] Siry explains. "But given the subjects we cover, we wanted the readers' experience of reading the magazine to be more like the experience of reading online."

So the whole magazine's gone sans serif, ragged-right, lots of gimmicky boxes on the pages...indeed, it is like reading online. Before too long, my eyes hurt and I just want to stop.

It's not the old *Wired* with its painful-to-the-eyes color combos, with a big dash of "if you can read this you have great eyesight" thrown in. Nope, it's just sad. Sadder yet: It comes from Time Inc., which used to take pride in its writing and design.

Masthead

Cites & Insights: Crawford at Large, Volume 7, Number 3, Whole Issue 87, ISSN 1534-0937, a journal of libraries, policy, technology and media, is written and produced by Walt Crawford, a senior analyst at OCLC.



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

Opinions herein may not represent those of OCLC or YBP Library Services.

Comments should be sent to waltcrawford@gmail.com. Comments specifically intended for publication should go to citesandinsights@gmail.com. *Cites & Insights: Crawford at Large* is copyright © 2007 by Walt Crawford. Some rights reserved.

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

URL: citesandinsights.info/civ7i3.pdf