Cites & Insights:
Crawford at Large

At Your Service

Marking CDs:
A Bad Idea?

A public librarian sent me a note almost a year ago (September 28, 2001), in response to a posting I'd made on PUBLIB regarding CD protection labels. I noted, "the graphic side of a CD is much more vulnerable to damage than the metallic/rainbow side. The information layer (pressed indentations with an aluminum coating for reflectivity) is just underneath a thin lacquer layer on the printed side. Scratch that side hard enough (or, more likely, 'identify' it with a chemically-reactive pen, since the lacquer side lacks the almost complete non-reactivity of polycarbonate) and you can disrupt the information layer and open the aluminum coating to corrosion."

The librarian sent this note: "We've been marking our CD Audiobooks with a permanent stamp pad ink (from Sanford; the bottle says it contains phenol), and tattletaping the case. We initially marked the discs on the clear hub with a permanent marker, but began using the stamp since it could include our address. Are typical 'permanent markers' considered corrosive? Does it sound like we should rethink how we've been marking these items?"]

I did a little quick research and concluded, "Yes, you probably should rethink how you mark CDs. Phenol is quite reactive and a solvent, and most of the pages I checked (and the relevant ISO/ANSI standards) caution against use of *any* solvent-containing pen or ink on the label (lacquer) side of a CD. Apparently some permanent markers are corrosive and some aren't. But phenol appears to be in the possibly-destructive category."

The Short Version: Don't Do It

In brief: *don't do it*. That is, don't use *any* ink or adhesive on the label side of a CD that isn't either certified for CD use or well known to be chemically

inert. And don't write on it with a firm-point pen. The lacquer layer is only 0.002mm thick—less than one ten-thousandth of an inch. (The polycarbonate layer is 1.2mm thick, six hundred times thicker than the lacquer—and polycarbonate is one of the most stable, non-reactive plastics.) On a pressed CD, that thin layer directly covers a sputtered aluminum layer. If the aluminum is exposed it will oxidize "within a few days," according to CD manufacturers.

Inside This Issue

Most permanent marking pens do use solvents; those solvents will almost certainly eat through the lacquer over time. There are some soft-tip pens that don't use solvents, but your best bet is probably to mark only within the clear hub area.

Fun Facts About CDs and CD-Rs

Most information you find about labeling CDs is only partly applicable—because it will be about CD-Rs and CD-RWs, not pressed CD-Rs. Some CD-Rs do use aluminum as a reflective layer; some use gold, much less subject to corrosion. But there's a dye layer below the aluminum (since CD-Rs can't use "indentations" as such—your burner doesn't have tiny little hammers), and that dye layer can also be corrupted if the lacquer doesn't hold up.

Some CD-Rs have writable surfaces. These are additional coatings on top of the lacquer, specifically providing a protected layer for a casual label. Still, it pays to be safe: use something with a relatively soft tip and don't press too hard.

If you're labeling CD-Rs, you'll probably use special self-adhesive labels that you prepare with an inkjet printer and apply with a special centering hub (e.g., the Stomper kit). For most use, that's a safe way to proceed—although even these labels aren't recommended for CD-Rs that you plan to keep for

many years or decades. That's not an issue for most of you, of course.

Most of this stuff came from the CD Information Center (CDIC) at www.cd-info.com/ CDIC/. That site has *enormous* resources. For even more information about recordable CDs, there's a *huge* "CD-Recordable FAQ" at www.cdrfaq.org. It makes fascinating reading and seems to be updated regularly.

Feedback: Your Insights

Jerry Kuntz of Ramapo Catskill Library System commented on my broadband coverage in the April issue:

> I suspect the unmentionable market for broadband is the file-sharing crowd, not video on demand. Beyond that, broadband makes a whole lot of sense if your home needs include a LAN where each client wants net access. For what it cost me for an extra phone line and a single 56K dieal-up account, I now get Optimum Online (Cablevision) and use a Linksys wireless router to network 3 machines in our house (me, wife, son).

> I would have predicted that broadband would fulfill both those markets quickly, then flatline. Who knows though; it certainly seems like the RIAA and MPAA are pursuing policy stances guaranteed to keep file sharing a thriving practice.

I pretty much agree with everything Jerry says, *including* the worth of broadband for many homes. As I responded:

> You have broadband because it meets your needs: Great! So do millions of other people, for good & understandable reasons. I don't have broadband yet, because it doesn't meet *my* needs, right now. But people at *PC Magazine*, "market analysts," and some in gummint think there's something wrong with me that I need to be Educated, and once I'm informed, I'll spend whatever sum is needed for next-generation broadband. There's the problem.

DisContent

Tracking the Ebook Niches

Donald Hawkins' thorough two-part discussion of electronic books (*Online,* 24:4 and 24:5) forms a fine baseline for considering this "major publishing revolution"—or, as I think of it, this set of potential niches. The growing range of products and services that can be called e-books still has much more hype than marketplace reality, but

it's an interesting (and confusing) aspect of repurposing content and attempting to separate content from carrier.

Counterpoints to Hawkins' Articles

The "advantages and disadvantages" chart in Hawkins' first installment omits the most serious disadvantage of all current electronic books: readability. But then, Gemstar seems to assert that print quality and typographic fidelity are irrelevant to readability. So would I, if I was pushing devices that lack both.

As far as I can tell, the "high-quality screens specially designed to improve the reading experience" on dedicated readers don't exist. The only reader with truly high-quality screens, the Everybook, was vaporware: it never appeared except as a prototype. Commercially available ebook readers are *at best* no better than typical notebook screens; reasonably priced ones are worse. Today, you can buy a notebook computer with 133dpi screen resolution; the best dedicated reader currently on the market tops out at 85dpi.

Hawkins also treats Microsoft's claims for Clear-Type uncritically, and I can only assume that he never tried out the download. Microsoft's claim of 300% improvement is nonsense, as far as I can tell. On a CRT, I found that ClearType made type *harder* to read by adding color fringes. The thought that Microsoft's involvement assures a big future for a field is no truer than the old belief that IBM always succeeded. Microsoft makes loads of mistakes, but the company is big enough to write them off.

In Part 2, Hawkins covers so much territory so well that a few minor problems were probably unavoidable. He calls NuvoMedia "NuovoMedia"—but the maker of the Rocket eBook is gone in any case. He devotes considerable space to the phantom of the marketplace, Everybook, which was probably doomed from the start as a device. Having called Stephen King's *Riding the Bullet* a short novel in the first installment, he calls it a short story in the second. It is neither: it is a novella.

It's hard to consider Gemstar's acquisition of NuvoMedia and SoftBook as a shakeout: usually, a shakeout requires a real market. There's a good reason that you don't see sales figures for dedicated ebook readers or ebooks in general, and it's not that the market is growing too fast to track!

Finally, I have qualms about the term "self-publishing" as applied to some of the e-publishing arrangements described. Some e-publishers seem to be little more than cut-rate vanity publishers; others fall somewhere in between self-publishing and vanity publishing. But that's a topic for another column.

Recent Developments

After this nitpicking, I should repeat that Hawkins did a fine job. His manuscript was completed in April 2000. By the standards of "Internet time," ebooks should be major success stories by now. With the partial exception of netLibrary and its pseudo-book model, that hasn't happened. Instead, the hope of a broader marketplace for specialized items and backlist titles seems to be fading, at least among the best-known ebook players.

Gemstar has had time to digest the only two serious manufacturers of dedicated readers. The company apparently found that Rocket and Softbook had such value as brand names that it dropped both of them. It now has the trademarked name Gemstar eBook and expects other companies to build the readers. Gemstar has also clarified its distribution plans: best sellers, and that's about it. Midlist? Not enough profit. A company spokesman says that backlists raise logistical issues for them, whatever that means. Obscure authors? No money.

The founder of NuvoMedia, Martin Eberhard, who *was* interested in broadening the publishing marketplace, offered this comment to *Industry Standard*: "Gemstar decided for its internal reasons that [ebooks] would be a good space to get into. It looked around and saw there were a bunch of players already making noise in the space, so it decided to buy them out and shoot them in the head."

The new reader marketplace isn't bookstores; it is (supposedly) places like Best Buy and Circuit City, just where avid readers always hang out. I've seen full-page ads for Gemstar eBooks—but so far, only in *TV Guide*, which makes them in-house promotions rather than paid advertising.

It would appear that Gemstar *must* maintain proprietary distribution technology; its only way of making money is through taking a chunk of each ebook sale. Gemstar says it wants to be the Yahoo! of reading. I see its model as the DivX of ebooks, and can only wish it the same success that Circuit City had with DivX.

Then there's MightyWords, the new name for "ematter" from Fatbrain. MightyWords appeared with great fanfare and a mission to offer midlength texts—longer than short stories, shorter than books—in PDF form and with a methodology that authors could live with. Authors prepared the texts and uploaded them in Word, RTF, or PDF form, and paid $1 per month per text for storage. Mighty-Words packaged the files into encrypted PDF to guard against casual copying, provided the Web site and some level of publicity, handled charge cards, billing, downloading—and retained half of whatever price the author set for each item.

MightyWords was off to a good start, but something went awry, right around the time Barnes & Noble acquired Fatbrain. The company notified three-quarters of its authors that it would no longer handle their works, and reduced the payment rates for most remaining authors. MightyWords has reduced its scope to the areas where it saw the likelihood of substantial sales—primarily business and technical literature. So much for novellas by new authors and all the other forms of medium-length fiction and nonfiction that are difficult to handle through traditional publishing. The content was there, but it wasn't profitable.

Finding the Niches

So far, the news is mostly discouraging. Not for book-lovers, to be sure, but for those of us who think there *are* niches that print publishing doesn't serve very well. Dedicated readers could have a substantial market for textbooks, where just reducing the weight that students must bear might justify the price of a reader. Some small e-publishers have been around for years, but they continue to suffer the same fate as most very small publishers, ignored by most reviewers and with sales generally too small to discuss. There may be one e-book with 6,000 sales over several years—but you need *dozens* selling in the high thousands before you have a marketplace.

Much of the publishing industry began because people were dedicated to books and willing to spend time and money toward relatively small returns. If ebook companies require quick bucks, it's unlikely that we'll ever see the promise of ebooks fulfilled.

> This "DisContent" column originally appeared in *EContent* 24:2 (April 2001), pp. 50-2.

Postscript and Update

The references to Don Hawkins' two *Online* articles may seem remote, but you can probably find them in full-text sources (or your print collection of *Online*, for that matter). I wrote this update in November 2000, a few months after my ninefold model of ebooks appeared in *American Libraries* and several months before Clifford Lynch gave us the right name for the least plausible part of the ebook market. Call them *appliances*, not ebook readers; that leaves the "reader" name open for software running on desktops, notebooks and PDAs.

I find it remarkable that some folks still believe in the ultimate triumph of ebook appliances; explore

eBookWeb to check out this particular straw man. Sure, there are new players from time to time—but most of them, like the pathetic little Franklin units, bundle in a bunch of other features. They're PDAs with big screens, not dedicated ebook appliances.

Gemstar stopped running full-page ads in *TV Guide*. The remaining partial-page ad, buried somewhere in channel listings, is nothing short of ludicrous: unchanging from week to week, month to month, it offers no way whatsoever to find out anything more about the Gemstar eBook. The Website still operates and claims best sellers—but Thomson/RCA stopped promoting the readers months ago and even Henry Yuen's unverified claims for appliance sales amount to roughly one percent of RCA's predictions. I suspect old RCA hands (if there are any) are remembering the Selectivision CED videodisc about now. (No, not laser videodiscs: this was a *contact* disc, like a super-high-density LP, and made just about as much sense as you might think. Its only real impact in the marketplace was to confuse people enough so that Pioneer's LaserDisc couldn't gain success. This all happened in the late 1970s and early 1980s. I heartily recommend *Current Technologies in the Library: An Informal Overview* by Walt Crawford (G.K. Hall, 1988).

MightyWords? It closed down January 12, 2002; the revised business model was too small to make any sense. As for netLibrary—well, you all know about netLibrary's travails, and thanks to OCLC it is still operating, albeit at higher prices. If you go back to my *American Libraries* article, you'll note that I thought it was "an intriguing model tuned to library realities." It still is.

Otherwise it's all rather sad. Sure, there are downloads by the hundreds of thousands, and print on demand continues to grow as an interesting counterforce to the top-heavy publishing industry. But the innovative models haven't worked out very well, and that's a shame.

The Filtering Follies

The CIPA trial is proceeding as this issue comes out. Meanwhile, Seth Finkelstein and others continue to research, think, and write about the problems of "censorware" (to use Finkelstein's term). Ssome interesting items I've seen since the last installment.

Anticensorware Reports

Seth Finkelstein (sethf.com) used to be chief programmer for the Censorware Project and now gener-

ates a stream of material at Seth Finkelstein's Anticensorware Investigations: http://anticensorware.com. He won an EFF Pioneer Award in 2001 for his work and claims to have examined more censorware programs internally than anyone else. He has an InfoThought mailing list; I haven't (yet) subscribed. You should be able to find these reports at either site mentioned. **Recommended.**

Last September, he wrote "BESS vs The Google Search Engine (Cache, Groups, Images)." I suspect Finkelstein pays particular attention to BESS from N2H2 both because it's marketed to corporate users and because it's regarded as one of the better (that is, less ideological) filters. This report considers the ways that BESS handles some of Google's most interesting features:

➢ Cached pages are banned entirely through the undefeatable "Loophole" category, probably because cached pages inherently evade server-based filtering. Finkelstein goes on to provide easy ways to undermine BESS's block, such as using a country-specific Google version or adding an extra parameter to the URL.

➢ Google Groups, a larger version of the DejaNews UseNet archives, is not *generally* blocked—interesting, since some Usenet groups have been hotbeds of textual licentiousness (e.g., alt.sex groups). Unless your BESS settings lock out Bulletin Boards, "with the correct search terms a person can read all the (text) porn they desire. Go figure." (You have to be clever; keyword filters stop obvious searches.)

➢ Google doesn't claim to do image recognition in its image feature; it just looks at the text name for the image file and text in the surrounding page. That's sensible; image recognition doesn't work very well. You would expect that BESS could, as a result, handle Google image searches as though they were text searches—but instead it lists all Google image searches as "Pornography."

As the report concludes:

It should be stressed that the dilemma here is intrinsic to censorware. Given a large undifferentiated archive, it'll either have to be banned entirely, or a reader may be able to retrieve items which would otherwise be banned. No matter which approach is taken by any censorware in a particular, the result is bound to be either too little or too much in terms of the concept of "filtering." There is no magic blocking program, only crude and ill-functioning attempts at control of information.

A later report, posted March 5, 2002, generalizes the image-search investigation: "BESS vs Image Search Engines." Here he concludes that the massive over-

blocking may be because image search engines cache thumbnails. (I'm charmed that he links to *RLG DigiNews* for an overview of the wildly varying results of image searches: it was a good overview, one I should have cited.)

What happens with the various engines? You already know that BESS calls all of Google Image Search pornography. It does the same for Lycos Multimedia Search, locking out *all* searches, not only image searches. BESS considers Ditto.com to be "Nudity Pornography Swimsuits" although Ditto.com itself claims to filter its results!

AltaVista? Unless you block Search as a category, BESS allows the searches—but it blocks all images from the AltaVista servers as Pornography, so you all you get are filenames!

Then there's FAST or AllTheWeb. FAST Multimedia Search is blocked as pornography—but if you use the general site (www.alltheweb.com/), you can search for nasty images to your heart's content. As usual, we see both massive overblocking and underblocking so substantial that it negates the whole point of a filter.

The Ethical Spectacle's Censorware section (www.spectacle.org) includes two interesting reports by Jonathan Wallace on N2H2 and BESS, with Finkelstein credited for his assistance in the second article. N2H2 is publicly traded, making its business success more open than privately-held companies—and, at least as of last summer, it wasn't doing all that well. You have to wonder about a press release heralding "strong third quarter financial results" when the quarter had $9.4 million in operating expenses and $2.1 million in revenue! The first report (dated August 8, 2001) is almost entirely financial; the second, dated September 1, 2001, discusses "N2H2's weak AI." Noting repeated claims of "sophisticated Artificial Intelligence" in preparing the BESS blocking lists, Wallace asked academics and others about current AI and tried to evaluate the reality of N2H2's claims. He concludes, based on real examples, that BESS lacks anything that could plausibly be called AI. There are other articles in the series, which I haven't examined. (Quick follow-up: a March 21, 2002 report in the Puget Sound *Business Journal* notes that N2H2, based in Seattle, has been delisted by Nasdaq for failing to have the minimum net tangible assets and stockholder equity for trading. It did *much* better in the first fiscal quarter of FY2002: a mere $1.6 million loss [revenues not stated in the news report].)

Responding to a Web4Lib note (from a UK academic librarian whose library uses Websense filtering), Finkelstein noted on January 16, 2002 that Websense locks out the Internet Archive Wayback Machine (www.archive.org) as a "proxy avoidance system." He goes into more detail in the March 13 report "The Pre-Slipped Slope—censorware vs the Wayback Machine web archive." After a quick description of the hundred-terabyte archive, he notes why it would be considered a filtering loophole and looks at how it's handled. BESS blacklists the site as a loophole (loopholes can't be disabled in BESS, according to Finkelstein). Websense uses a different name but does the same thing. SmartFilter 2.0 blacklists the Web archive in every category; the situation with the current version is murkier.

This report includes a longer version of Finkelstein's claim that censorware "is about control, not filtering. The goal of censorware is to construct an escape-proof blinder-box for what a person is allowed to read." He goes on to demonstrate the "slippery slope" that results—which, as he says, is more like taking a flying leap off a sharp cliff in the case of most filters. The history of the Web: that's all pornography! Looking for a picture of the Statue of Liberty? You dirty old man—filters will put a stop to that. After all, if you're allowed to look at *any* pictures or read *any* archived material, you might find a way to get to the hot stuff—and we can't have that!

In a brief item on March 20, Irene Graham of libertus.net (http://libertus.net), an Australian site on freedom of speech and censorship, reported that SmartFilter blacklisted the entire site as "Sex." Finkelstein checked: the site was indeed blocked under that entirely inappropriate heading—but then, libertus.net discusses the problems with filtering software. "After this is publicized, SmartFilter'll eventually change it, and the censorware blacklist will be assumed perfect once again…"

Other Items

Last December, I noted and recommended Benjamin Edelman's expert testimony in one of the CIPA cases. Even though it was redacted (portions blacked out), it was still fascinating, worthwhile background material. He's now posted a redacted "Expert Rebuttal Report" following expert reports submitted on behalf of the government. (I can't find the pro-filtering reports; I haven't found much recent pro-filtering research. Maybe I'm not looking in the right places.) If you read the earlier testimony, I **recommend** you read the followup.

Edelman demolishes most of the arguments of the pro-filtering experts, showing how and why most filters will both underblock and overblock. He also notes significant flaws in the pro-government testing methodology. But then, even the government expert's claim for effective blocking of questionable

sites doesn't say they're perfect. In one case, the government expert claimed estimates of anywhere from 82 to 96%—but Edelman, going through a deeper list of sites, finds blocking effectiveness (for "free adult sex" as a Google search term, eliminating two clearly non-offensive sites out of 797 total) running anywhere from 73% to 89%.

A remarkable and obvious flaw in one opposing expert's testimony comes in analyzing updates to filtering program site lists. When he looked at updates, he tested *only* sites that he originally considered wrongly blocked—*not* new sites that might now be incorrectly blocked! Naturally, this means that filter accuracy always improves: if you only count improvements, what other result could there be? Even by the other expert's numbers, the best claim is that Websense is "only" blocking about 6.7% of legitimate sites and BESS is "only" blocking about 6.9%--but there's reason to believe that the real figure may be closer to 8%.

Why worry? So one out of every twelve legitimate sites can't be viewed. So what? At least the filters will keep out four out of five nasty sites, or maybe as many as nine out of ten! If there are 100,000 prurient sites on the Web, kiddies will only be exposed to 10,000 of them: problem solved!

Mind if I remove every twelfth chapter from the books in your library? I can't tell you which ones: that's not how filtering works.

It's almost refreshing to turn to a different set of issues as expressed by Brian Smith in "Why I'm against mandatory porn filters," the January 4, 2002 issue of *Ex Libris* (marylaine.com/exlibris/xlib126. html). It's short and **recommended**, making an important point: overblocking and underblocking aside, mandatory filtering inherently blocks *protected speech* in a venue where the library has already provided open access. "The core of the constitutional argument against mandatory filtering doesn't depend on the shortcomings of current filtering technology, but on the right of access to a designated public forum." Looking back through this series, I don't see any mention of "filteReality"—Brian Smith's extensive site on public libraries and Internet filters. The site is www.filtereality.net/ and doesn't get updated all that often (the most recent update is listed as 18 December 2001), but it does include a strong list of links not only to anti-filtering sites but to pro-filter sites as well, along with some other good sections. **Recommended**.

An odd item comes from a librarian at Boca Raton Public Library. While evaluating CyberSitter, she found that the library's *home page* was blocked because the teen/Young Adult link includes the word "adult." Keyword filtering also wreaked havoc on the

library catalog and subscribed databases; look for *Naked Lunch* and you'll get back titles containing "lunch." ("Naked," after all, is a Nasty Word and must be blocked at all turns.) This librarian also found that sites previously reported as inappropriately blocked were still being blocked.

Odder and more difficult: "Filtering software: the religious connection" by Nancy Willard, posted on February 24, 2002 at Responsible Netizen (netizen.uoregon.edu/documents/religious2.html). This long, footnoted discussion considers connections between major filtering companies and conservative religious ISPs and religious filters. I'm not entirely convinced that the connections she makes are conclusive—that the use of filtering software by the religious right automatically means that the filtering software *as generally released* has a religious bias. This is a case where I would **recommend with caveats:** read it carefully.

What about the Internet Content Rating Association? A March 22, 2002 *Wired News* report notes the release of "ICRAfilter," a user-controlled filter that works based on ICRA labels, used by "more than 50,000 websites" in the year since the system was introduced. The article quotes Karen Schneider, who probably knows as much about filtering as any librarian; she called ICRA "misguided in a sincere sort of way" for a variety of stated reasons. Schneider's stance (and mine) is that what you do with your family's computer is your business—but somehow people always want to impose filters on library and school computers. The ICRA spokesman says the group "takes no position" on filtering in schools and libraries—but he thinks ICRAfilter should be *everywhere*, whether it's activated or not. What does ICRAfilter do when *any* filtering is active and it hits unlabeled sites?

Trends and Quick Takes
To Blog or Not to Blog

OK, I'm guilty: I wrote an article about Weblogs as part of a cluster of *American Libraries* articles on the circle of gifts, and I rely on a dozen or so Weblogs to point to items for commentary in *Cites & Insights*. On the other hand, I don't do a Weblog—and almost all the Weblogs I check regularly are atypical, according to the Blogging stories I've been seeing lately. That is, a majority of Weblogs appear to be online diaries of a sort; most of those I check are focused sets of library-related links, sometimes annotated, rather than extended mirrors for the creators. I have no idea what Blake Carver (or other contributors) ate for breakfast on

March 12, but LISNews almost always points me to one or two worthwhile sources each week.

I was reminded of that distinction—that most Weblogs are much more personal (and self-oriented) than the ones I monitor—by a charming *Wired News* piece by Farhad Manjoo, posted February 18, 2002: "Blah, blah, blah and blog." Manjoo notes the strongest indication that Weblogs are now mainstream: NPR ran a piece on them. And there have been stories all over the place. This piece says that Weblogs have now crossed a "tipping point" (one of those memes that I haven't picked up on) with Evan Williams of Blogger saying there are "a million different kinds of weblogs." A later estimate is that there may be half a million Weblogs in all, so Williams' comment on variety may be hyperbolic.

You may know some of the backlash. John Dvorak made fun of most bloggers looking for ego gratification; two bloggers gave out the second annual "Anti-Bloggies" for blogs that are boring, lame, obsessed or weird. (I've looked at the Anti-Bloggies—but I'd never go back to one of them unless someone paid me.)

Here's what I found peculiar about the *Wired News* piece: comments from Dave Winer. Somehow, he seems to think that *everyone* should be building Weblogs—that they are social goods of some sort. He's not the only one. "Asked if he'd like to live in a world where virtually everyone blogs, Williams chuckled and said, 'Yeah, I think it would be a great thing. It's not that you want to read them. But people have the desire to express themselves, and I think it's tremendously powerful activity. If you write everyday, your writing improves, your thinking improves.'" I'm not sure I can buy that as a general proposition—and I am sure that most good writing is something more than spur of the moment jottings.

What would you do with 150,000 library-related or librarian-related Weblogs? (There are at least that many librarians just in the United States.) Already, I find that I'll only add about one out of three new possibilities to my "try it for a while" list; with even a thousand bloggers in the field, the few good new ones might be buried in the noise.

Tunes to Rent

I'm willing to pay a monthly subscription for DVDs, when the process works as well as Netflix does. After all, there are precious few movies I want to see more than once or twice. Music's a different issue, and I wonder whether subscription music services make any sense at all. So does *PC Magazine*, I believe, reading the March 12, 2002 "Pay-and-play music services" writeup carefully. EMusic gets an Editors'

Choice for its unlimited downloading of 128K MP3 tracks that can be burned to CD-Rs for $10 a month (on a year subscription), but 128K is (as the magazine *correctly* says for once) *near* CD quality and most of the music comes from independent labels—you won't find most fave raves. Still, EMusic's terms make sense for the consumer.

Not so for Press*Play* (love that orthography!) and RealOne Music. The first costs $10 to $25 a month; that gets you between 30 and 100 downloads (of 128K WMA files, which should sound a little better than 128K MP3) and 300 to 1000 streamed plays—and the right to burn between 10 and a grand *20* cuts to CD-R. You can only burn two songs from any one artist per month; essentially, you're paying for the right to prepare one mediocre-sounding CD a month and listen to a few other tracks. RealOne is worse: no downloads, no ability to copy to a PDA or MP3 player, and—get this—downloads aren't cumulative: you can only have a hundred tracks on your PC at any one time, for which you pay $10 a month. Such a bargain!

The Short Century

BBC spent something like $4 million in 1985 to compile a vast record of the state of Britain. They called it the Domesday Project in honor of the 1086 inventory. The new inventory would be "the mother of all time capsules." The results were stored on two 12" laser discs—discs that could only be viewed using BBC Micro computers.

You know what's coming, even if you haven't read the story in *The Observer* (March 3, 2002) or elsewhere. The original Domesday Book is in fine condition after twelve centuries. The new version is essentially unreadable. "Few [of the BBC Micro computers] were purchased, and only a handful are left in existence." There's a project to rescue the data. It might eventually work.

The Good Stuff

Seltzer, Larry Jay, "Password crackers," *PC Magazine* 21:3 (February 12, 2002), pp. 68-71.

"Beware: This is a disturbing story. Many of your illusions about the security of your systems and data will be destroyed." It's in the gearhead "Solutions" section of *PC Magazine*—and it's about *legitimate* programs that crack PC passwords at various levels. They're marketed to help administrators recover lost passwords—but, of course, anyone can be the administrator of a small company!

It's a worthwhile story, with tips on constructing hard-to-crack passwords and which applications are more or less open to attack. There are also important secondary tips—for example, passwords don't do much good if intruders can get their hands on the PC being "protected." Also, long passwords that mix numbers and letters are effective against brute-force attacks and dictionary-based attacks. But you knew that already, didn't you?

Isaacson, David, "Instant information gratification," *American Libraries* 33:2 (February 2002), p. 39.

"Many of us—even librarians who should know better—have come to expect our information needs to be gratified as instantly as our sensual needs. We are too quickly satisfied." Isaacson writes a fine perspective in something like 800 words. If you were too hurried to read the full-page "On my mind," go look it up. He's not attacking the Web, but he is raising pointed issues, among them the difference between instant facts and meaningful information.

McInerney, Claire, Alex Daley and Kay E. Vandergrift, "Broadening our reach: LIS education for undergraduates," *American Libraries* 33:2 (February 2002), pp. 40-3.

This article deserves thoughtful reading and consideration. The authors are all in Rutgers' School of Communication, Information, and Library Studies; they describe a number of undergraduate programs being carried out by library schools. The programs don't lead to degrees in library science—which, given professional requirements, would be fairly useless—but to such degrees as Information Management and Technology or Information Systems. Graduates probably won't become librarians—but they may wind up in many other jobs for which library-related skills are worthwhile. The advantage for the library schools is quite clear: they become more central to their institutions, and their faculty gain more varied teaching experience and student contacts.

Steers, Kirk, "Rev up your net connection," *PC World* 20:3 (March 2002), pp. 103-14.

Yawn. Yet another collection of obvious and arcane tips. Maybe so—but I found this set of "29 ways to keep your cable, DSL, satellite, or dial-up modem link cruising" refreshing. Partly that's because of the final choice: some of us may *choose* to keep using dial-up modems for a few days more. Quite a few of the tips can be useful for dial-up connections as well as broadband. It's also less cutesy than some tip roundups and seems blissfully free of

"memorize this four-key combination and avoid pulling down a menu!" pointers.

Sauer, Jeff, "The matrix," *EMedia* 15:1 (January 2002), pp. 24-33.

Noting that *EMedia* speaks primarily to professionals in the "emedia" fields (CD, DVD, and related fields), this lengthy discussion of DVD authoring tools may still be worth your time. That's particularly true in two specific situations:
➢ You're planning to do video editing and create DVDs
➢ You believe that inexpensive digital videocams, video editing, and DVD burners will turn us all into filmmakers.

In the first case, the article discusses failures in the industry, where things stand now, and some of the astonishing variety of tools ranging from $50 to $120,000. In the second case, you get a hint of the real issue: The tools don't make the artist. I could do without some of the snideness—"In reality, many consumers struggle to assemble simple photo albums from piles of photo lab-processed prints"—but the underlying issue is nonetheless true.

The tools don't make the artist. I suspect you'll see me expound on that theme elsewhere at some point, so perhaps I should clarify a bit here. Yes, word processing makes it feasible for me to do as much writing as I do while also working full time, being lazy, reading, and watching TV. Yes, newer materials have allowed artists and sculptors to achieve effects previously impossible—and, for some, lowering the initial bar makes all the difference. My brother turns out to be a capable composer; score generation using a Mac (and extra time after retirement) made it possible for him to show that creativity.

But in no case do the tools provide the creativity. Given the best camera in the world, I will not be a great photographer: I don't have the eye to judge composition and lighting, although I can certainly appreciate the work of a good photographer. Given the best word processing system, someone who can't maintain a coherent thought process won't produce good columns or great novels. Film-making involves a whole set of related skills and creative talents. Computers and digicams can lower the bar; they can't, and won't ever, turn us all into videographers.

Carroll, Sean, "Top 100 undiscovered Web sites," *PC Magazine* 21:4 (February 26, 2002), pp. 86-97.

I'm not sure this is "the good stuff," but I've given *PC Magazine* a well-deserved raspberry for their lists of top hundred Web sites, noting that they're all business to a laughable degree—as though

".com" is the only domain on the Web. So it's worth noting a modest turnaround. The sites here are peculiar in many cases, but they're also interesting and cover a slightly broader range. And there are non-coms: www.annoyances.org, www.linuxdoc.org, www.cdc.gov, www.archive.org, www.spaceflight.nasa.gov, a "chemsoc.org" URL that's too long to repeat, www.nrp.org, www.pbs.org, us.councilexchanges.org, www.time.gov, www.poets.org, www.poynter.org, and pbskids.org. That's more than one out of eight: a *lot* better than one out of a hundred. The list also includes some interesting sites I'd never heard of: worth a look.

Fritz, Mark, "DeCSS down but not out," *EMedia* 15:2 (February 2002), pp. 8-9.

This doesn't belong here for two reasons: it should be part of a copyright cluster and it's too brief to mention. It does belong here for one key reason, pleasantly surprising for an industry-oriented magazine.* Read the first sentence: "DeCSS, a simple program meant to allow Linux users to view DVD movies on their personal computers, has become the focal point of several controversial and acrimonious legal battles." Consider the phrase between the two commas.

DeCSS is not the product of some pirate factory. It was not designed to help people steal copyright material or copy DVDs. It was written by a Linux user because no vendor had released DVD playback software for Linux. The user wanted to watch *legally purchased* DVDs.

Mark Fritz says that right up front, before delving into the various legal quandaries. Would that other technology journalists (and journalists in general) were as thoughtful in framing their coverage. For that matter, the brief news piece is a good summary of what's happened so far.

(*OK, that's misleading. *EMedia* comes from Online Inc., which also used to publish *Online*. I've been dealing with and writing for them for years, currently in *EContent*, and they always meet or exceed my expectations for ethical and journalistic standards. So this isn't a great revelation—but it's still remarkable for a casual news item.)

Besides, you'll want to pick up this issue for the next item:

Bohannon, WK, "Footloose and PC-free," *EMedia* 15:2 (February 2002), pp. 24-9.

If you have to give presentations on the road, you may need to haul a portable projector. They've been getting lighter and better—and this article considers the next step. Maybe you don't need to carry your notebook computer! Some of today's units ac-

cept PC Cards and can run PowerPoint presentations directly. If you don't need to change things at the last minute, that may cut your total weight in half. It's a good, thorough article from a man who claims to have given the same presentation, "using the same slides, about 1,000 times."

Of course, you could use my personal trick—speaking without PowerPoint or other AV—but that doesn't work for some presentations, and many speakers blanch at the thought of working without a net. If you're one of them, read the article.

Ginsburg, Isaac, "The disregard syndrome: a menace to honest science?" *TheScientist* 15:24 (December 10, 2001), followed up by Garfield, Eugene, "Demand citation vigilance" and several letters in 16:2 (January 21, 2002). All downloaded from www.the-scientist.com.

Ginsburg's opinion piece posits that too many scientists are writing articles without doing exhaustive literature searches—and that referees don't catch the problem and prevent duplication of already-published results. Thus we get redundant publication and even more additions to the overload of journal articles. Eugene Garfield calls it the "disregard syndrome"; Robert K. Merton calls it "citation amnesia." The opinion piece suggests why it happens and offers proposals to help avoid it. The letters all support the view that there's a growing problem; one of them offers the slightly utopian suggestion that everybody should just scan in all the "pre-Medline" articles out there and put them on the Internet, so that scientists could find all the articles so much more easily. Garfield's follow-up commentary suggests a "science court" and signed pledges that authors have searched all appropriate databases. I **recommend** the original opinion piece as something worth thinking about—and I'm reminded once again of why I'll never be a scholar!

Grotta, Sally Wiener, "Focus on photo editing," *PC Magazine* 21:5 (March 12, 2002), pp. 72-6.

You never know where you'll encounter excellent articles. I usually regard *PC Magazine*'s "Solutions" section as geek central, as befits the subtitle "Tools & tips for the Internet age." But this one's a worthwhile introduction to today's most important photo-editing techniques. Grotta writes well. **Recommended** if you have reason to do photo editing and aren't quite sure how to get started.

"20 years of technology," *PC Magazine* 21:5 (March 12, 2002), pp. 98-165.

I could review this blockbuster as four parts or fourteen, but it all works together. *PC Magazine* be-

gan publishing in March 1982; I've read every issue. It's a shadow of its 400-page self, but it still appears 22 times a year and offers the most thorough reviews and special features of any personal computing magazine.

The first section, "Technology in America" (pp. 98-126) consists of a seven-page introductory essay by a writer I generally find annoying followed by a series of two-page area-specific essays (e.g., Office, Home, Communicatons). I could nitpick, particularly about the introduction (which quotes a group of extreme early adopters to show how we're all connected all the time and loving it—like quoting the Pope on universal religious attitudes), but most of the writers are good and it's an interesting set of views. Read it skeptically, but it's worth reading.

No skepticism needed for the next section, "Living history" (pp. 137-59), which recounts the history of the magazine and its field in a series of pages covering one to four years each. Michael J. Miller wrote the whole thing (or that's what the byline says) and did a fine job. **Recommended**—as are the final two two-page spreads. The first (pp. 160-1) offers tales from PC Magazine Labs over the years; it's the most important testing facility in the PC field (and the only one to uncover universal "gaming" of video-card tests—every single maker tried to cheat the lab's tests at one time or another!). The last (pp. 164-5), "We told you so," is a series of quotes from the magazine including some hits—and some misses. Consider, for example, this letter in December 1983: "We believe the arrival of the PC's little brother [PCjr] is as significant and lasting a development in the history of computing as IBM's initial foray into microcomputing has proven to be." Or this from April 12, 1988: "The 'OS/2 decade' has begun." They missed a memorable quote around the time the first PC/AT appeared, when one writer said that the 286 provided more power than any user really needed—but it would make a great server.

Kinsley, Michael, "Social hypochondria," *Slate,* February 28, 2002 (slate.msn.com).

Maybe I'm including this as a form of special pleading, after boring you all with the literacy story. The National Center on Addiction and Substance Abuse put out a study "noting with alarm that a quarter of all the alcohol sold in America is consumed by teen-agers." Later, the New York *Times* noted the study was wrong because it "had not applied the standard statistical techniques in deriving that number." More deep manipulation problems such as those in literacy reports? Not really: the survey sample was 40% teenagers and the reported numbers didn't correct for that gross over-

representation. The actual figure is more like 11%, which may or may not be a cause for concern.

Kinsley notes, correctly, that any competent news organization should have questioned the standalone number in any case, regardless of sampling problems. Let's say there's a report that people over 50 are tax cheats—they account for 25% of all tax fraud cases! (I'm making this up.) But then you learn that 40% of all taxpayers are over 50, which puts a slightly different spin on the story.

I would assert—and Kinsley implies—that NCASA's original report *deliberately* overstated the situation in an effort to create a crisis. Is it obviously alarming that teenagers consume *any* alcohol? That depends. Would this headline, "Study Concludes Teenagers Drink Less than Half as Much as General Population," raise cries of alarm? Probably not.

You've seen deliberate overstatement before. Anti-gambling groups alerting us to the debilitating effects of gambling almost always cite the *handle* of casinos and racetracks rather than the *take*, and most reporters don't understand the difference, even though that difference can range from 10:1 to 50:1 *or higher* depending on the situation. A quick refresher: the "handle" is the total amount wagered; the "take" is the amount that stays with the casino or race track—not the profit, but total wagers minus total wins. If I play a typical Nevada quarter slot poker machine for an hour, the *handle* for that machine will probably be $60 or more (that's four hands a minute at one quarter each, a leisurely pace)—but the take is *extremely* unlikely to be more than $3 to $6, unless I'm getting truly awful cards. Moderately skillful non-counting blackjack and craps players will wager *at least* 50 times as much as they lose: both games have 98% payout or better for good bets. If that wasn't true, most casinos would be empty, as the gamblers would all be bankrupt.

Kinsley is more interested in the problem of "social hypochondria," our tendency to exaggerate the perils of the social ill of the moment and our silly belief that those ills can be entirely corrected. "America is not, as it sometimes seems, a society lurching from one acute social crisis to the next. It is a basically healthy society with lots of chronic problems that exist simultaneously, can and should be ameliorated, but will never go away." He goes on to explain the forces that make it hard for people to see that simple truth: politics, media, and lawyers.

"According to a recent study by the respected National Center for Credulity and Alarm, Americans are twice as likely to swallow a phony statistic about a social issue, and almost 2.7 times more likely to find it alarming, as citizens of either the European Union or the former Soviet bloc."

"Google today," downloaded March 1, 2002 (www.google.com/corporate/today.htm).

I never thought I'd be citing a corporate puff piece—but then I never thought I'd be citing lengthy law review articles either. Maybe this discussion resonated with me because I've always been offended by the assertion that the business of business is to make money—usually strengthened to "the *only* business of business is profit" or some paraphrase of that.

I believe that *great* businesses, whether small or large, begin by doing something that matters and then proceed to find ways to make that profitable. Garbage collection matters, concrete mixing and delivery matters, serving really good pizza matters. I never worked for HP, but I believed in the HP Way.

Maybe that's why I find this statement so interesting. The heart is "10 things Google has found to be true," from "Focus on the user and all else will follow" to "Great just isn't good enough."

Block, Marylaine, "Telling people what libraries do" and "Who's going to preserve zine content?", *Ex Libris* 134 (March 8-15, 2002) and 135 (March 22-29, 2002). (marylaine.com/exlibris/)

Short notes on two good, brief commentaries. The first concerns the many services that libraries offer and that people don't know about. The second hits even closer to home and concerns the "permanence" of e-zines (focusing on library-related zines) and the likelihood that they could disappear, archives and all, if their creators lose interest or are no longer able to foot the bills. She's kind enough to mention *Cites & Insights* along with *Library Juice, NewBreed Librarian*, and some others. In the latter case, she suggests that someone like Wilson or Ebsco should negotiate to archive the zines and index it along with other library-related full-text publications; in the former, she doesn't really have answers but thinks we need to look for them.

Text-e Part III:
Enter the Mandarins

Reading without writing, new architectures of information, authors and authority: the final segments of the five-month text-e "virtual symposium." As before, I reviewed the English version of each paper and the predominantly-English commentaries as a single packet—and reviewed the three packets in a single evening. I also reviewed the moderators' conclusions—but, as I begin this segment, not the two extra weeks of commentaries on those conclusions.

Sad to say, not one of these segments excited or informed me, at least not positively. The tenth segment informed me that "my kind" really had no business in text-e: it was really for the mandarins, specifically the European mandarins, to commune together and assure themselves of their superiority, as with so many invitational symposia. Think of these comments as ignorant heckling from one of the great unwashed, one of those bourgeois fools who fill the Web with trash (defined as "anything other than peer-reviewed scientific and technical papers") and fail to recognize that their betters deserve obeisance, not dispute. I am no scholar; I lack tenure; I have no business among these giants. Mea culpa, mea maxima culpa.

But, having mistakenly walked two-thirds of the way across the chamber of the mandarins, far be it from me to turn and flee at this point.

Dan Sperber:
Reading without Writing

Sperber's thesis in this paper is "that the revolution in information and communication technology may soon turn writing into a relic of the past: it will be replaced by the automatic transcription of speech—whereas reading is here to stay." Which he immediately follows with a copout: "My aim, however, is not to prophesize, but to reflect on the future with the help of tools developed within the cognitive and social sciences."

He suggests (correctly) that his view is less radical than the suggestion of William Crossman and other that "talking computers will finally make it possible for us to replace all written language with spoken language" and that, this being possible, it *will happen*—and be a good thing! I've seen that suggestion and regard it as hopelessly dystopian. Sperber's future strikes me as equally unlikely—and I fail to see how his "tools" make his case.

I'm struck by how people know everything about daily life in prehistoric societies or more recent ones. Sperber says, "In most of the human societies that have ever existed, children become competent adult [sic] without the help of any formal teaching." Is that true? "Formal teaching" does not (to me) imply classrooms and grades; it does imply a regular (formalized) routine by which one generation passes along crucial knowledge to the next generation, usually with certain people filling the role of teacher. I would have (naively) thought that formal-

ized intergenerational training was almost part of the definition of "society."

Next we're told, "writing and reading are acquired, if at all, through a lengthy and intensive process of deliberate training in interaction with a teacher." At least half of that is demonstrably false. I began reading before entering school; so do *many* children whose parents read to them. There was no "intensive process of deliberate training" and no teacher was present; there were parents and siblings who cared about the stories in magazines and books and shared them with me.

I could pick out a fair number of flaws and questionable assertion. He says, "The greater the number of people who read and write, the greater the benefits involved in being able to do so oneself"—but scribes could earn good livings when they were the only writers, livings no longer available.

Today's speech recognition software doesn't work well enough to yield faster overall output than that of a skilled typist. Sperber's answer? "I take it for granted, however, that these shortcomings will be overcome…" with no consideration of the limits that the English language (at least) may place on potential transcription accuracy. It's the same handwave that "futurists" always use for technological limitations. Technology solves all problems: just wait a year or two.

Along the way, Sperber provides enough background to suggest that his thesis is irrelevant, unless he really means to say that people will cease *being able* to write—which makes no sense in a future where we still read. Maybe you use a calculator most of the time, but surely you know enough arithmetic to calculate a tip in a restaurant or check the addition on a bill that shocks you—we have not, in fact, become wholly innumerate. Even if you dictate 90% of your writing, why would you stop being able to write at all? And, for that matter, what would it mean, if you were still editing the results to achieve the quality of written work?

Sperber says "speech is several times faster than hand writing or even typing." According to my dictionary, "several" means "more than two and fewer than many," in which case Sperber is wrong. I'm far from a professional typist, but my error-corrected rate is between 70 and 80 words per minute. Other than advertising stuntman, I know of nobody who speaks much faster than 160 words per minute, and certainly nobody who achieves a standing rate of three times my mediocre typing speed (that is, 210 to 240 words per minute).

Even if Sperber's fallacious assertion was true, so what? If I could create a rough draft at my full typing speed, that would be 4,200 words an hour—a book-length draft in five four-hour days. Some low-end romance novelists and porn writers may achieve throughput like that, but for mere mortals a thousand-word hour is highly satisfactory output. Being *able* to dictate, say, 9,000 words an hour would make no difference.

"The creative potential of writing does not come from the movements of the hand but from those of the eye." I always thought that the work of the *mind* was far more important than the eye or the hand—and there are certainly blind writers of some note.

Consider these three fragments: "The few exceptions—olograph wills and scented love letters for instance"; "When you read, you loose the extra input"; "the teaching of writing would rapidly loose much of its significance." What do they have in common? Illiteracy or incompetent editing—much as though Sperber had used the tools he foresees as everyone's future. If that's true, it suggests a danger in abandoning writing: those who do not write their own text may not be effective at editing their own writing. (Sure this is nit-picking—but Sperber's paper is one of the invited papers in this international formal symposium, and as such should be edited not only for clarity but for grammar and spelling.)

Part of me says it wouldn't matter if Sperber's "likely" future came to pass; another part says that it's improbable. My major reaction is that the paper is unconvincing and badly in need of editing. **No recommendation.** His closing paragraph may deserve quotation in full:

> One can imagine anything. On the other hand, to speculate in a manner that is both informed and reasoned is difficult. Difficult but not altogether impossible, I hope.

Commentaries on Sperber

David Prater from Australia notes the benefit of speech-to-text conversion for people who suffer from typing-caused RSI—but, of course, even today's speech recognition software can and should be used by RSI sufferers. He then says, "Typing would have to be one of the most unnatural and damaging activities brought about by the computer revolution." I've been typing a *lot* longer than I've been using a computer keyboard, and I can guarantee that a Royal manual will tear up the wrists faster than a Microsoft Natural keyboard. I thought typewriters came into common use several decades before computers. Did I miss something? Prater does apply a reality check: "The only problem seems to me to be that we have been hearing about advances in voice recognition for at least a decade now." Actually, voice recognition *has* improved considerably in the

last decade; it's now plausible as the starting point for writing, particularly if you're not a touch typist or you do have RSI. But, of course, that's *not* the "only problem"—he goes on to note that people can't use the technology unless they have access to it. To which Sperber responds that if we don't bother to teach people to write, we can afford to give them all speech recognition computers. Oh, and for poorer countries, we'll skip reading as well, using computers for everything! Right...

Stevan Harnad does what Steven Harnad does. He speaks from Mt. Olympus—we learned in earlier discussions that Harnad knows better than RLG, LC, OCLC, and CLIR combined regarding the need for digital preservation efforts—but begins with a mere four-page comment on the nine-page paper, including only three of his other articles as citations (and, one citation where he's not listed as an author!). Does he make it through the comment without touting "skywriting"? What a silly question! Harnad may be right in some of his comments on Sperber's paper, but before he finishes he drops back into his same old groove. In one of Harnad's many comments, he introduces "mouthwriting" as well.

The moderators conclude that Sperber's "provocative" paper "has given rise to a highly constructive debate." I missed some French and Italian comments, but what I read struck me as neither constructive nor very interesting. I've seen more constructive debate and progress in a day's thread on a single topic on Web4Lib, albeit with many fewer words and generally without such wonderfulness as "apodictic" (which does, I suppose, have a slightly different meaning than "dogmatic," although Harnad's writing is both). I learn that one non-English comment included the "forward-looking" suggestion that "children should be taught how to type at school." Am I wrong to suppose that most children *do* learn how to type in school, at least when they're using computers? I seem to remember a typing class in 1960; I never knew Modesto was so many decades ahead of society as a whole or that Americans learn to type while Europeans do not. The moderators conclude:

> Real conclusions, of course, are pending; they might be another few years, perhaps decades to come. But the discussion over the past two weeks between Dan Sperber and participants in text-e signals a very real need to think concretely about the issues surrounding the emergence of speech-to-text technologies about the transformations we are witnessing and those we are participating in. The format of text-e, at once oral and written, mediate and immediate, is helping us make headway.

I have no coherent written comment to offer, and I must have missed the oral portion of text-e.

Stefana Broadbent and Francesco Cara:
New Architectures of Information

Broadbent & Cara are at IconMedialab and begin by informing us that, during the past four years, "we have carried out hundreds of observations of people using the Web." They claim to have used "every *possible* type of Website, wapsite, portal and digital product." [Emphasis added.] They go on to say, "the downfall of the economy has come as no real surprise. The first generation of Websites has been built with a radical misunderstanding and misrepresentation of users' expectations and practices."

Is IconMedialab Europe's equivalent to Jakob Nielsen's absolute assurance that he knows everything about how every Web site should be constructed? I haven't visited their own site to see if they adopt the children's-book approach Nielsen does, but they have the same omniscience and willingness to contradict everyone else that makes Jakob so resistible.

"*All* Internet users can be divided into three main categories: expert, naïve and what we call 'light' users." Any such division means that nothing sensible can be said about any category. Which they proceed to prove by telling us that Experts "remember Mosaic" and "download software" and that naïve users "still think that Internet is magic and that there is only one site for each topic." *Everyone else* is a light user: "They use it one or two hours a week, doing a bit of email, looking up practical content, such as train timetables and cinemas." Light users represent a "relatively new user population, which has emerged over the last couple of years." Do you sense that a few million Internet users, at least in the United States, fall into none of those categories?

If not, maybe the next paragraph will help.

> Most light users have very stereotypical behaviors: after six months of usage of the Internet they stop even trying to do searches through a search engine and consult systematically the same six or seven sites. Search engines are too complex and deliver too many results to weed through. These users usually give up searching because it is too costly for the results they obtain.... Their only way to discover new sites is to collect URL addresses in the press or other traditional media or from friends and family.

At this point, I conclude that at least one of three cases must be true:

- American Internet users are so vastly different than European users that no sensible discussion can encompass both groups.
- This document did not survive translation to English in good form.
- Broadbent & Cara are so prone to grotesque oversimplification and, indeed, "misrepresentation" that everything in the paper must be viewed through an unusually skeptical lens.

The word I wrote immediately below that paragraph was highly impolite, and I won't repeat it here. It is the same word I wrote as a four-letter summary of the entire paper once I'd plowed through the remainder. I found the paper to be a complete waste of time, and suggest that you skip to the next section if you value your own time.

Broadbent & Cara are almost sneering in their attitude towards the benighted Light Users. They may "stray from their sites using the links on the page of their favorite sites, but just like toddlers who never stray more than a few meters from their mothers, these users don't venture too far away from their familiar sites." They "don't read full pages of text." And, repeated again, "most Internet users don't use search engines, only visit a few sites—and always the same ones…"

Then they tell us why (after some blather about slowing adoption of the Internet). Web sites are *too hard*: the Web is a "very complex technology" that requires "learning, motivation, and especially proof of its value." We don't explore because the "cognitive cost" is too high.

Later, the claim that none but experts ever read full Web pages is stated more strongly: "users essentially skim the pages and don't read *any* text" [emphasis added as we devolve from not reading full pages to reading nothing at all].

There's more—a lot more—but I found it either impenetrable or absurd. N**o recommendation.**

Here's the final paragraph—without comment, because I can't think of anything remotely sensible to say, particularly about the second sentence (and am not sure when the singular "medium" ceased to be part of the language):

People do not read on the Web as they do on paper—but why should they? It would be like using a Ferrari to mow the lawn. The Internet is a powerful, still largely underused interactive media that can allow both editors and users to create new forms of access and new ways of sharing content. Its relational capacity, of which hypertext is a first instance, is what makes this media unique. But to this day, the relational potential of the Web has only been approximated. When deployed more fully, it will indeed change the way we relate to text and multimedia content.

Commentaries on Broadbent & Cara

Slim pickings this fortnight, at least in English, but how do you work from a starting point like that? Richard Minsky observes that the worldwide Internet user population is *currently* growing at a 15% annual rate and notes that "very few sectors" in the real world show that kind of growth. In a much longer message, he takes Cara to task for blaming the "written book model" for stalling the widespread adoption of the Web." He thinks that overstates the case (which is putting it mildly), discusses economics, and goes into an odd discussion of browser oligopoly and the difficulty of writing fancy Javascripts in Notepad—but who does that? He goes on at length, including the claim that "soon we will have out-of-the-box AI database driven web design software." I always love claims of artificial intelligence!

In yet another comment, Minsky takes text-e to task for needing too many Back button steps to get where you want to go. The site navigation was changed; I still find it clunky.

A moderator chimes in making comparisons between text-e itself and the paper; Minsky (again!) responds with a Kids These Days commentary that also brings in "nanotechnology, molecular computing, artificial intelligence, VR, robotics, cellular integrated chipsets" as things that will allow new kinds of interaction. He seems to think that in "less than 30 years we may find that there is an internet being programmed by AI robots to communicate with each other for the purposes of their own evolution (if they are not already doing that!)" Naturally, "AI robots" delight Minsky but he fears that "fears of the humans…may slow it down, and may even destroy civilization as we know it." We learn that "greed and lust" are "central to the development of the www."

Maybe there were first-rate, cogent, worthwhile commentaries in French and Italian. Maybe there *are* hordes of "AI robots" out there. I'm just an old Luddite (in Minsky's eyes, at least, as far as I can tell), so how would I know?

The bad news: this was pointless in my view. The good news: the portions were small.

Umberto Eco:
Authors and Authority

Once again, instead of a paper we have a conversation—an interview by Gloria Origgi, one of the moderators. Between the exigencies of the interview format and translation, the obvious course of action (to avoid seeming to criticize Umberto Eco) is to say

that what's here may not represent what Eco was actually trying to say—or that I'm insufficiently literate and sophisticated to get his set of points. All of which may be true.

Eco says that filtering is "the fundamental problem of the Web. The whole of the history of culture has consisted in the establishment of filters." Eco seems to be saying that culture does *not* arise from creativity; culture is defined by what it rejects.

At the second question, I hit an "arggh" point: the question struck me as ludicrous—and the answer as equally difficult.

> G.O.: But today we have authoritative filtering systems that belong to this very means of communication, that is, search engines.

> U.E.: That's not a filtering system. There are already polemics on the fact that search engines "filter" only information that has been paid for. I don't believe in the possibility of automating the filter's function. The only solution is that there appear authorities, external or internal to the Web, that constantly monitor what is found… There should be specialized monitoring groups, for example the International Society for Philosophy could…monitor all the philosophy sites… Now if I trust the International Society for Philosophy, which tells me: "This site on Kant is rubbish," then I won't use that site…"

I'm on Eco's side in refuting the suggestion that a search engine is an "authoritative filtering system." I do not understand why Eco feels the need to slander *all* search engines (including Google) with the false claim that they filter "only information that has been paid for." The transcription doesn't show "some" or "defective." It's an absolute statement, absolutely wrong for more than one search engine.

Remember the subtitle on this installment of text-e commentaries? *Enter the mandarins*: We must have authorities to determine what is worthwhile on the Web—constantly monitoring, constantly providing thumbs up or thumbs down. The keepers of high culture must prevent the rest of us plebeians from becoming lost in the rubbish that dominates the Web. Eco has the good sense to recognize that Web users might have difficulty recognizing that a site is in fact an Authority—but he's convinced that such authorities are both necessary and possible.

Eco correctly says that filtering authority—where "filtering" means "judging sites," not "blocking sites" as in software filters—does not constitute censorship but a form of consultation. Such consultation takes place on the Web all the time, but it's done through patterns of links and almost certainly does *not* satisfy Eco's needs. To me, there's a difference between an "authoritative" site and trusting that links on that site will represent good material,

and an Authority with the role of labeling sites good or bad. The first arises through organizations and reputation; the latter requires some form of certification—an authority of authorities?

Eco regards *samizdat* (self-publishing) as "the other problem with Internet" as those fine filters, the publishers, are left out of the loop. "At least it relieves publishing houses of a load of useless manuscripts…" For some reason, Eco believes that, without publishers, there will be no worthwhile criticism of Web-based manuscripts: "the function of orientation provided by criticism is gone." Why should that be? Critics don't work for publishers (do they?).

Eco regards it as "extremely dangerous" that the canon might be weakened. He indulges in a KTD argument: "new generations are born with a mechanism for attention that is adapted to the screen." A bit later, though, Eco points out the additive nature of new media and technology: "It has never happened in the history of humanity that the introduction of a technological means killed off all the practices of the previous means." My motto, "The new tends to complement the old," is far less literary and uses a mere seven words when 25 are readily available. Eco believes "the increase in information…will not have an effect on the use of books; on the contrary it will increase it."

Then it gets strange again, or maybe things have gone quite differently in Europe. Eco had gone to an "important Italian publisher" with his plans for a CD-ROM product five or six years ago. The publisher "said first that there was no future for CD-ROMs, and second, that the winning formula would not be a CD but something or other that Philips had brought out…" So far, so good—but "nowadays this important publisher makes CD-ROMs by the bucket." Except as inclusions in computer-related books, I'm not aware of too many contemporary publishers making "CD-ROMs by the bucket"; the medium seems in a permanent state of decline.

Now that I've trashed Eco on the issue of authority and filtering, let me say that he offers thoughtful and convincing comments on copyright and royalties and on the promise of print on demand. But he says that Sperber's prediction (above) is "highly likely" and uses as a parallel that the "new generations that use calculators will lose" the capacity to do mental arithmetic.

He offers a page on libraries that I am unable to make sense of—but from what I can glean, it appears that the notion of public libraries is either foreign or irrelevant to him. The closing question is about the "original" of a written text in an age of word processing—and, as Eco notes, that concept is pretty much gone. He offers a fascinating example, if it's true: an

Italian writer offered $5,000 for the original manuscript of a novel, who then had the printed book typed up, then went through with a pen to give it the appearance of an edited manuscript!

No recommendation.

Commentaries on Eco

The three English-language "debates" I printed total 26 pages, roughly three times the length of Eco's interview. Fifteen of those pages come from Steven Harnad. One begins with the thundering, all-caps title: "PEER REVIEW IS AND ALWAYS WAS THE FILTER AND 'AUTHORITY'." Can you guess how Harnad feels about anything other than peer-reviewed scholarly articles?

> Faced with the heterogeneous hash on the Web today, most of it trash, even if 'filtered' by Google's link-count economy, it is understandable why people might think that filtering it all is a new problem, unique to the web, and THE problem. But it is not; it is none of these things: the problem is old, predates the web, and solved.

He goes on (he *always* goes on) to tell us that "for science and scholarship" the filter is peer review, with the journal name being the "tag." He introduces "skywriting" again shortly thereafter, and another endorsement of the Web: "this vast vanity press—this global graffiti board for trivial pursuit…" The solution: "If one is serious about one's inquiries, one will restrict them to the texts of qualified experts, certified by their qualified fellow-experts."

If you're interested in areas where peer review doesn't exist (or covers only a minority of the worthwhile material), you're not serious. And, of course, only "qualified" experts can "qualify" other experts. If peer review had only been around for longer, and if we had all had Harnad's absolute confidence that only peer-reviewed material is worth pursuing, we would have been spared much grief along the way—we would *know* that the sun revolves around a fixed Earth, for example.

It is Harnad's special skill that, having made a point, he finds it useful to quote as many passages as possible from the original paper, giving his "no need" response to each one—thus turning a one-paragraph commentary into five pages. Along the way, he manages to draw out some of Eco's condemnation of popular taste that I missed in the original interview—although Harnad's contempt for anything that isn't a peer-reviewed scholarly article, and for those who read such trash, far exceeds Eco's. He throws in another argument for Harnad's Optimal and Inevitable Virtual Book and snidely suggests that watching TV is probably killing off "a lot of actual and potential literacy," also slandering news-

papers along the way. His list of references is back to his usual: all five named references are from Harnad, S., once joined by others.

David Klemke makes a terrible, unforgivable mistake:

> To suggest that to solve this whole problem of "unsolicited" information by just using such a thing as "Peer Review" is quite ludicrous. Although, yes it will filter some sites and render a small part of the web "reliable," it will still leave the vast majority of this etherial library left unfiltered.

I presume Klemke writes English as a second or third language, and will ignore the difficulties with the paragraph itself. The mistake, of course, is to suggest that a Harnad statement could be ludicrous. Harnad…um…gently corrects Klemke and tells us what's on the Web, in lieu of the entirety of peer-reviewed articles: "a lot of idle chatter." Or, later, "hot air." And, since Harnad explicitly ignores "art," he says, "there is no such problem to solve… The way to filter [chatter] is to ignore it." He adds a PS against cries of elitism—but the PS is unsatisfactory.

Quite apart from the fact that most of us, most of the time, are not looking to the Web for information or ideas readily available in the refereed STM literature, Harnad's universal solution ignores the need for interpretation and popularization. Harnad says, "think of an afflicted relative, and what sort of information you want your physicians to rely upon." That's not the question. Most of us cannot run to a physician any time we have a symptom or other puzzling situation; to do so would be to cause the complete collapse of already-overburdened health care systems. We look for guidance as to whether we *should* seek out a doctor or whether our symptoms might better be treated with pseudoephedrine or a couple of aspirin. I doubt very much that I could approach Medline with symptomatic information and retrieve the appropriate peer-reviewed articles. I doubt *even more* that I could understand those articles sufficiently to draw reasonable conclusions, and I suspect there are millions who are even less capable than I am of translating medical journal articles into everyday information. (Guess what? If I want to understand why electrical outlets require grounding, and what to do about that, I doubt that scholarly articles on electronic momentum—or whatever—will be of much use to me. Even electrical engineering articles would give me far too much information in a form that would be nearly useless.)

I will assert that, for most Web users looking for useful information (rather than art, literature, and ideas—another category altogether), refereed scholarly articles are generally the *wrong* things to retrieve. The right things to retrieve work at another level, a

mediated level that brings expertise in a more understandable form and applies scholarly results to the human condition. Harnad's solution does precisely *nothing* for those users, and is as a result no solution at all. To write off all use of the Web except Experts Speaking to Experts as being graffiti, hot air, and chatter is offensive and absurd.

Eco is *primarily* concerned with areas such as philosophy and literature—areas where Certified Experts have even less plausible claims that the world should ignore anything that doesn't receive their blessings.

Another "debate" begins with a full-page paragraph from Roberto Casati that seems to say (I think!) that the link structure of the Web is the answer to universal peer review—thus, I suppose, that the top Google results are *inherently* authoritative. Minsky goes a large step forward: "The lack of peer review may be the internet's greatest contribution to intellectual and artistic freedom…common sense might prevail" in the long run. "Perhaps the ivory tower is not the best model for creative evolution… 'Scholars' build on each other's misconceptions without examining the underlying axioms." I find Harnad's "if it ain't peer reviewed, it's rubbish" attitude horrifying—but Minsky's delight at the notion of discarding peer review entirely strikes me as equally horrendous. Harnad returns to suggest that "a welcome solution" would be a Web consisting "of all and only this canonical Gutenberg corpus." As far as I can tell, the *only* new material that Harnad considers even possibly worth considering is "non-published teaching materials." Otherwise—Harnad has spoken, and if it doesn't come from the Ivory Tower, it should be ignored by reasonable readers. "The rest is just dictascript, which need no more be 'navigated' than what transpires on the airwaves of chat TV or a hairdresser parlour."

Then there's "Let a thousand flowers bloom," with Stephen Downes' assertion that "the internet is replete with filters" and that "there is much less of a need for filters than Eco supposes. Why? "In a literate society, the vast majority of people are able to distinguish between quality work and that of the less-qualified fringe. People are moreover able to critically assess even the authorities and draw their own conclusions. Filters are perhaps necessary for a pre-literate culture, but we have progressed well beyond those medieval days."

I would *love* to be as sanguine as Downes but I am all too aware of how many people believe in astrology and creationism, are convinced that bottles of "remedies" containing less than one atom *per bottle* of some ingredient can cure diseases related to that ingredient, regard global warming as nonsense

cooked up by socialist one-world crackpots, appear to believe that the Bible was dictated to King James' staff, in English, directly by God—people who believe that software filters can in fact eliminate all "pornographic" material on the Web while leaving intact all "worthwhile" material. While not as ready as Harnad to write off all that does not come from Certified Experts, I am unwilling to assert that the "vast majority" of us can, *in every field*, distinguish not only between well-written and badly written work but between correct and misleading work.

Unlike Eco, I believe that we have many "filters" that are *legitimately* not Certified Authorities but have built up trust through other means. I regard *Current Cites* as a generally reliable set of pointers to worthwhile literature (although not a perfect one, and certainly not intended to be complete); I regard Charles W. Bailey's Scholarly Electronic Publishing Bibliography Weblog similarly, based on experience in both cases. Neither of those are true filters in Eco's sense, particularly since neither attempts to label worthless articles or sources in their areas. And perhaps that's a problem: for a variety of reasons, negative filters tend to be few and far between. That may be human nature. I know that "Cheap Shots and Commentary" interests me less than "The Good Stuff," and that I would never consider attempting to identify all the articles in my fields of interest that I thought were rubbish. It's hard on the soul to be consistently negative; it's harder on professional relationships.

This segment offended me the most because of its appeal to authority and some participants' contempt for humanity in general. And yet, and yet… **No recommendation.**

Moderators' Conclusions…

The moderators offered up a seven-page "conclusions and questions" document on March 14, inviting another two-week "concluding discussion." As I write, that discussion is underway. I may have additional comments when it closes. I find the paragraph of "numbers" interesting in some odd ways:

> Some 3,500 people went to see the site every two weeks to read a new text. Over 60,000 people visited our site, 795 people subscribed in order to be able to participate in the debates, and 2,292 people downloaded eBooks—a total of 7,277 individual eBooks. The archived debates contain 662 comments and questions sent by the invited participants, the speakers and all those who took part in text-e.

Hmm. On average, those who downloaded eBook versions of the papers (none of these were long

enough to qualify as "books" in any real-world sense) did so for roughly 3.2 papers each, and the "average paper" received 728 downloads. Putting it another way, only one out of 26 visitors chose to download even one paper.

"Out of the 40 invited participants, 15 never posted a comment." Thus, of the elite—those "properly" in the symposium—almost one-third opted out of any active role.

I'm less convinced than the moderators that there have been no "virtual symposia" on the Web in the past or that text-e has been the Web equivalent of a formal symposium. The moderators are convinced that text-e *improves* on "real" symposia. I may not be qualified to comment; while I've participated in quite a few conferences, I am not among the elite invited to symposia.

We learn that moderators edited commentaries before they were posted. I find this astonishing, given the results—but then, the primary language of the source was French, and the French versions of all comments might be crisp, grammatical, and sensible. "We also received at times one long, fifty-line paragraph with no caesura, interesting in content but too much of a strain for the eye." Yet I *read* paragraphs that seemed to go on forever, some of them at least 40 print lines long!

"With each debate, we deepened our knowledge of the problems." I found that, by the end of the event, I lost sight of what "the problems" might be.

…and My Own, If Any

"Would that I had missed text-e altogether." That's the first and easiest conclusion: too many pages of *Cites & Insights*, too much of my time has gone to this commentary and the associated reading. Even as part of the non-peer-reviewed rubbish on the Web, I might have made better use of that space.

I found Roger Chartier's article enlightening and thought-provoking. Although I found most of Harnad's commentaries as annoying and offensive as much of his other work, his paper was also effective within its clearly restricted sphere. Jason Epstein did a fine job, although he may have overstated his case. That's three out of ten—and that's a decent percentage for a journal, but not for a symposium!

It could have been worse—both the papers and the commentaries. It could have been a *lot* better, but that's true of life in general. Stevan Harnad could learn to make a point once, and to allow for the possibility that he's neither omniscient nor always right—and pigs could learn to fly.

Postscript

After writing the above and before unsuccessfully attempting to cut it in half, I printed off the brief biographies of the invited contributors and the list of invited symposium participants. Unfortunately, text-e is one of those sites that doesn't want you to print some portions: the bios print on a deep-gray background (on a laser printer) and the text runs off the right edge of the page. It would be worse with A4 paper, which is slightly narrower than standard American letter size.

Who are these people? Europeans except for Jason Epstein, an experienced American editor and publisher. Roberto Casati and Dan Sperber are both at CNRS, the Institut Nicod (cosponsor of text-e). Two biographical paragraphs include superlatives: Theodore Zeldin is "recognized as one of the most important and original historians of our time" and Jason Epstein is "one of the most brilliant editors of the twentieth century." Looking at the list of invited participants, I find no names that I recognize except Régis Debray, but that is a comment on my provincialism; these must be the elite of French and Italian scholarship (with one Brit, the "other Michael Gorman," thrown in).

I conclude that text-e was intended to be Europeans speaking to Europeans, the intellectual elite determining the future amongst themselves. They will ignore the bits of opinion thrown at them by this American monkey, and perhaps that's as it should be.

The Details

Cites & Insights comes from Mountain View, California. Magazine editions are those received here; seasonal and other references are from a California perspective.

URL://cites.boisestate.edu/civ2i6.pdf