



Kerry Webb

kerry.webb@alianet.alia.org.au

Hot issues, Goth librarians and web solutions

Journal prices

There's been discussion for a long time on journal prices and the way they're subsidised by contributors and their institutions. Ted Bergstrom from UCSB has been a strong critic of publishers' activities, and his site at <http://www.econ.ucsb.edu/%7Etedb/Journals/jpricing.html> has a number of papers and discussions on the subject. He also has a tool at <http://www.journalprices.com> that can search for major journals and rank them according to their price per article or citation.

Trusted repositories

RLG and the US National Records and Archives Administration have produced a draft Audit Checklist for Certifying Digital Repositories at http://www.rlg.org/en/page.php?page_id=20769. It's designed to be used by staff working in digital repositories (to assist in their planning) and for people considering depositing their materials (to know what to expect from a repository).

Cost-conscious digital preservation

One project worth keeping an eye on is *espida* at the University of Glasgow (<http://www.gla.ac.uk/espida/>). They're addressing the issue of developing a model of the relationships, roles and responsibilities, costs, benefits and risks that are relevant to institutional digital preservation. In doing this, they'll be integrating digital preservation with strategic thinking, planning and culture at the university.

Digital repositories event

With guest speakers from Cornell University, MIT Libraries Digital Library Research Group, HP Labs and HarvestRoad, Open Repositories 2006 will be worth attending. It's a series of linked events to be held at the University of Sydney from 31 January to 3 February 2006.

The first will be the Digital Repositories Forum 'The Well Integrated Repository' (http://www.apsr.edu.au/pen_Repositories_2006/forum.htm), to be followed by the workshop 'Managing Openness in Digital Repositories' (http://www.apsr.edu.au/Open_Repositories_2006/Openness.htm). There'll also be the DSpace User Group Meeting 'Building Communities with DSpace' (http://www.apsr.edu.au/Open_Repositories_2006/DSpace06.htm).

Neologisms

As one who watches warily for the acceptance of new words (our very own Macca dictionary is notorious for letting just about anything in) I was interested to see what Merriam-Webster is including in its latest offering. There's a sample

of 'New Words and Senses' from the *New 2005 Collegiate Dictionary* at http://www.m-w.com/info/new_words.htm. The list includes *chick flick*, *metadata* and the new sense of *neoconservative*. There's also *cybrarian* — which is probably going out of style by now.

Black arts

Sarah Houghton is a Goth, so she says. And that explains why she called her site 'Librarian in Black' (<http://librarianinblack.typepad.com>). She set it up as a one-stop shop for tech librarians, which I suppose covers the readership of this column. Hmm. Maybe I should ask her if she's interested in some sort of merger — as long as I don't have to wear black.

Google Print — the saga continues

You will have noticed that I've featured a number of stories over the past few months about the Google Print project and the various objections to their plans. Well, it would be quite easy to be reporting each move and counter-move for the next year or so, but surely there must be a better way of keeping up to date with the affair. And so there is. Charles W Bailey Jr, famous for his *Scholarly Electronic Publishing Bibliography*, has collected together a number of articles and news reports at <http://www.escholarlypub.com/digitalkoans/category/bibliographies/>. The bibliography also includes a number of comments, some of which offer additional information on the topic, as well as opinions.

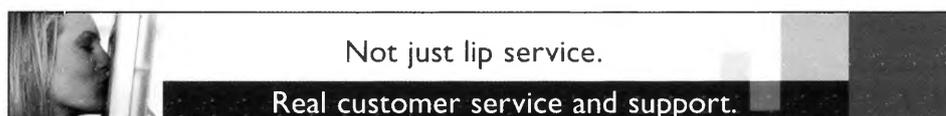
Design problems and how to avoid them

There's a very good article by Kim Dudley called 'Six design problems' at <http://www.communitymx.com/content/article.cfm?cid=F4476> that should be required reading by all students of web design. The problems cover fonts, graphics, layout and writing for the intended audience, and he provides some examples, tips on how to avoid the problems and links to useful articles.

What, really, does AAA mean?

One of my favourite questions is 'what are we trying to achieve?' If you can answer that one, you're well on the way to arriving at your destination, or at least you can be sure that you're heading in the right direction. Which is why I like the article by Dan Zambonini about AAA Accessibility. He points out why virtually no websites achieve the AAA level — no matter what they claim. OK — you could say that he's being deliberately provocative and that it's the spirit of accessibility that we should be aiming for, but I think he's right in saying that so many people are making these claims (or ask-

'Six design problems' should be required reading for all students of web design...



ing for them in tender documents) when they don't really know what they're talking about. See the article at <http://www.onlamp.com/pub/wlg/7730/>.

When a printout is needed...

Some of us are happy reading things online; others like to print them out — especially recipes from websites, which often come with various printing formats, including index cards. But in such cases, any print format is better than none. In a short article at <http://www.searchengineguide.com/welford/005897.html>, Barry Welford discusses some of the options for creating printer-friendly content, and endorses the use of style sheets. He makes a good point that you should have some idea of how many of your users will want to print your pages, but it's probably not too hard to make that sort of assessment.

E-mail problems

There's a warning for all of us in a recent e-mail problem. A local cookery school has a monthly newsletter that they send out in PDF format, at around 400KB for each issue. It's a useful little publication and I always take note of their cooking tips and news about upcoming courses. A couple of weeks ago, something happened, and they started spitting out another copy of the newsletter every few minutes. Now, most of us — whether at work or at home — operate on a mail system with certain limits on the inbox. In a case like this, with several megabytes of mail coming in every hour, you reach the limit very fast. It took several hours to cut off the flow of e-mails, by which time most subscribers would have done what they could to blacklist the sender, never to be accepted again. The school posted an apology when it was all cleared up (blaming the problem on a system that hadn't kept up with their Windows software patches) but by then how many people would have noticed, I wonder. It would have been a better idea to put the newsletter PDF online and just posted a link. But now it's too late.

How to do it

Of course, it's easy to put a newsletter online if you have your own website, FTP software etc. But what if you don't? Well, that's where you use someone else's system: YouSendit at <http://www.yousendit.com>. All you do is specify who you want to send it to, what file is to be sent, and hit the 'send' button. The file is then uploaded from your computer to their site. There's a size limit of 1GB on each file, and they hold the file for 7 days or 25 downloads. It displays ads on the site when you go to download the file, but surely that's a small price to pay.

How much is enough?

There has been regular discussion in recent years about how much information has been created in recent years, usually in comparison with the total amount created to date — and

that's not even considering how much of this information is any use at all. One study (based on the year 2002) that's gone some way towards answering these questions has been published by Berkeley University at <http://www.sims.berkeley.edu/research/projects/how-much-info-2003/>. It breaks it up by storage media (print, film, magnetic and optical) and by information flows (telephone, radio and TV, and the internet). It's not the final word, but it's a brave effort.

Where's it all going?

In case you're wondering where all the web traffic is going, Alexa has produced a collection of data from millions of their users who have toolbars on their computers. There aren't many real surprises in the top three — Yahoo!, Microsoft Network and Google. After that, it really starts to get interesting. Check it out at http://www.alexa.com/site/ds/top_sites?ts_mode=global?&none#.

New internet journal

The Internet Society has announced the *IETF Journal*, a new publication produced in cooperation with the Internet Engineering Task Force. It aims to provide an easily understandable overview of what's happening in the world of internet standards with a particular focus on the activities of the IETF Working Groups. Each issue will highlight some of the hot issues being discussed in IETF meetings and in the IETF mailing lists. See more at <http://ietfjournal.isoc.org/>.

Where have you been today?

I've written occasionally about the use of breadcrumbs on web pages, and there's always plenty of discussion about whether they're worth the trouble. Following a recent article by Jared Spool at <http://www.ue.com/brainsparks/2005/09/26/value-of-breadcrumbs/>, the ensuing discussion brought up an interesting side issue related to copyright. It also introduces the quaint concept of teleportation around websites.

New trends in document classification

Automatic document classification, the holy grail of information specialists, continues to spawn new concepts (and usually new acronyms). A recent one is Latent Semantic Indexing (LSI), which examines all the words in all the documents of a selection and calculates similarity measurements for each document or for individual terms — sounds familiar, doesn't it? There's more explanation at <http://www.searchengineguide.com/marshall/005944.html>.

And who said you can't take it with you?

They're still earning, even though they're dead. *Forbes* magazine has published a ranking of the Top-Earning Dead Celebrities — from Elvis, to Marilyn Monroe, to Dr Suess and John Lennon. It's live at http://www.forbes.com/home/lists/2005/10/25/highest-earning-dead-celebrities_deadceleb05_land.html. ■

Save yourself the trouble of typing URLs!

This column (with URL links) can be found on ALIANet at: <http://alia.org.au/publishing/incite/>.

Contributions and suggestions for this column are always welcome. Please contact Kerry Webb via e-mail kerry.webb@alianet.alia.org.au.

Softlink web based library management software.

www.softlink.com.au or call 1800 777 037



Automatic document classification — the holy grail of information specialists...