Kerry Webb

kerry.webb@alianet.alia.org.au

Librarians and
Google share the
same mission: it's just
the way we go about
it that's different...

Will digitisation and free text searching kill metadata?

You can't keep a good Wombat down

The Stumpers-L list (also known, for complicated reasons, as the Wombats) is one of the great gems of the net. It was hosted for many years at Dominican University in the US, but that association unfortunately came to an end in December. The good news is that the list has been given a new home under the auspices of Project Gutenberg, under its new name of Project Wombat. See the new, improved version at http://www.project-wombat.org/.

But we've only just come to grips with 1.0

We always have to be wary of overhyping and buzzwords, so I've noted the use of the term 'Web 2.0' for a while without bothering to delve into its true meaning. That's no great loss, as the mix of technologies and concepts known by that name is taking a while to make its presence felt. But when you want to know what it's all about, a good place to start is Tim O'Reilly's excellent paper at http://www.oreillynet. com/pub/a/oreilly/tim/news/2005/09/30/ what-is-web-20.html.

Google recognises us

A couple of months ago, I mentioned the Librarian in Black site. It continues to be a great read. In December, the LiB mentioned Google's 'Newsletter for Librarians', which also looks like something worth subscribing to. The first issue — at http://www.google.com/ newsletter/librarian/librarian_2005_12/ newsletter.html — starts with the assertion that librarians and Google share the same mission: to organise the world's information and make it universally accessible and useful. No doubt about that: it's just the way we go about it that's different. You can register to receive the newsletter by e-mail or read it through Google Groups. The archives (including a very useful article by Karen Schneider on evaluating the trustworthiness of sites) are at http://www.google.com/services/ librarian_newsletter.html.

More on digitising books

HarperCollins announced in December that it plans to convert around 20 000 books in its catalogue into digital form, to try to inhibit copyright violations on the internet. It's doing this at the same time as it joins in the flurry of legal activity against Google's efforts to scan copyrighted books in libraries.

And just to keep focused on who's who

It's not easy to remember which is which in this volatile world, so Danny Sullivan at Search Engine watch has provided a summary of which bit of Google is doing what with which bits of the library/publishing world. It's at http://blog.searchenginewatch.com/ blog/051107-233134/ and there's a link to the subscribers-only section of his site that has all the latest news on this hot topic.

Not just cards

Gary Price in Search Engine Watch has written a short article on some of the innovative ways that libraries are tailoring their online catalogues to the needs of their current users. Examples include book reviews and cover images, lists derived from the content of the books in their collections, RSS feeds and different types of user interfaces. See it at http://blog.searchenginewatch.com/ blog/051108-202655/.

Finding audio stuff

I've not quite got into the podcasting revolution — partly because I don't have a pod, but also because I can't find time in my busy life to download the content and listen to it. Anyway, when I do start podding, I'll check out Podzinger at http://www.podzinger.com/. The site takes podcast feeds, runs voice recognition over them to produce a full text version of each podcast, and then provides a full-text search capability. At first, I thought that they'd be publishing the 'transcripts' of the podcasts, but instead they link to the home sites for those. The voice recognition seems to only be used to extract the keywords for the search index. It's an interesting concept.

More on electronic resources

Coming out of the Engineering Libraries Division of UC Davis is a Punch List of Best Practices for Electronic Resources. Their task force examined the field and came up with 11 major issues and listed the best practices to address them. Apart from copyright and digital quality, others on the list include perpetual access, co-branding and access authentication. See the paper at http://eld.lib.ucdavis.edu/ punchlist/ and you'll even discover what a Punch List is.
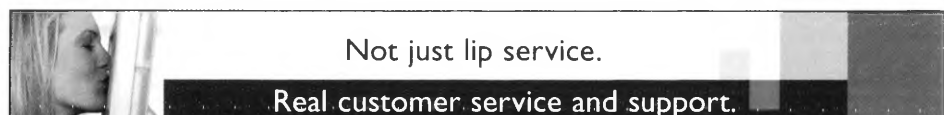
All about document delivery

DocDel.net is a directory for document delivery services and users. Their finder services lists subject specialist, full-service providers and suppliers outside the USA. It also features about a dozen articles on various aspects of document delivery and links to a number of discussion lists relevant to the subject. See it at http://www.docdel.net.

Now what was that tune?

In the quiet period over the Christmas break, I found out about Song Tapper and wasted a

while playing with it. It's based on the perennial problem of the tune that you know but aren't sure of its name. There are various ways of working this out, but this one is based on the rhythm of the song. What you do is tap out the tune by hitting the space bar (they recommend that you hum along as you do it, to get the rhythm correct). But does it work? Well, there are only a few thousand songs in the database, so as long as you select one of these there's a chance that it will find it. I tried 'God save the Queen' and it suggested something called 'My country 'tis of thee' as the first choice, but the correct answer came next. Similarly, it recognised 'Home on the range', but when I tried a few others, well — I guess I can't have been tapping tunefully. Try it yourself at http://www.songtapper.com/.

Out of the vaults
Record company Universal has announced that more than 100 000 tracks from records not currently available will soon be released for downloading. Among the artists involved in this release are Jacques Brel, Nana Mouskouri and Brigitte Bardot. Universal say they realised they could profitably sell a relatively small number of copies of a song as long as a compact disc did not have to be manufactured and distributed. You can track the progress of the initiative at http://new.umusic.com, but be warned — it's quite flashy.

Travelling bags
The Monash Public Library Service is encouraging their patrons to travel far and wide ... and take their library bags with them. Go to http://www.monlib.vic.gov.au/libraryBagImgs/AroundtheWorld/index.html to see some of the places they've been.

And, talking of old stuff ...
It's the sort of thing that doesn't happen often, but when it does you're pleased that someone had a bit of foresight. If you need an older version of a piece of software, go to http://www.oldversion.com/. There you'll find, for instance, Eudora from 7.0.1 all the way back to 1.4.4, together with a summary of where the major new features came in. Similarly, there are many of the previous versions of Internet Explorer, Opera and Firefox. Another useful site that specialises in archived browsers is http://browsers.evolt.org/. One thing you'll notice about all of these is the way that the software has grown in size over the years — a system that might have taken less than a megabyte when it started is now usually well over 10 MB, and that's a long download on a dial-up connection.

Firefox add-ons
One of the nice features about Firefox is the developer community that's doing its best to enhance your browsing pleasure. They've created some neat little features that don't re-quire much effort to install. One that I like is Greasemonkey, which allows you to load Javascript code to a web page that you've accessed. You could, for instance, block all the ads from your Google search results (although if enough people do this, the advertisers aren't going to want to use Google, and then where would we be?) For more information about this, including a list of the scripts that you can download, see http://greasemonkey.mozdev.org.

The end of metadata?
From the US federal government, comes the suggestion that with the spectacular developments in free-text search engines, maybe they don't need information management standards. An RFI issued late last year asks if the advances in search technology have made metadata unnecessary. It's a fair question, even for those of us who believe that metadata (if properly assigned and exploited) can be a great help in finding the information that we need. But the point is that, outside of a few little pockets of enthusiasm, it hasn't been used very much at all. The question needs to be asked. See the story at http://www.fcw.com/article91114-10-17-05-Print/.

... perhaps not
But all is not lost. One expert on search engine optimisation and marketing points out the value of limited use of META tags on websites, saying that the Description tag in particular has 'soared in importance'. Her column is at http://www.fcw.com/article91114-10-17-05-Print/.

They're never through playing games
The National Sport Information Centre, a part of the Australian Sports Commission, has launched its catalogue on the net. The catalogue (at http://www.ausport.gov.au/nsic/catalogue.asp) provides access to one of the best sports collections in the world, including 20 000 book titles with a subject focus on all aspects of sport, 2000 current and ceased sports journals and 5000 videos or dvds covering coaching, training and competitions related to sports played in Australia. It also covers journal articles, electronic resources and newspaper clippings and files.

The big clean-up
There have been a few stories on the news recently about New Orleans and its recovery from hurricane Katrina. There isn't much to celebrate — even the prospect of thousands of American librarians gathering there in July is probably too distant to be sufficiently cheery. Meanwhile, they get by as best they can. The New Orleans Public Library's website has some information about their recovery efforts following hurricanes Katrina and Rita, and especially their fund-raising. See it at http://nutrias.org. ∎

With the spectacular developments in free-text search engines, maybe we don't need information management standards. Is metadata unnecessary?