



Kerry Webb

kerry.webb@alianet.alia.org.au

In the story at <http://news.bbc.co.uk/2/hi/technology/3897583.stm> there is some good news, and a couple of quotes that — taken out of context — might cause you to stop and think twice. The main story is about a group of robotics researchers at the University Jaume I in Spain who are working on creating service robots for libraries. The group suggests that one potential application would be for a robot to 'fetch the book and, as directed by the web user, turn to the correct pages and scan the text and images.' Maybe. There is also a reference to some complementary work being done in Japan. As for the quotes, how about 'A library is a semi-structured environment,' or 'Libraries are the perfect environment for robots'.

And what does Hollywood think of libraries?

In the movie *I, Robot* there's a scene where the bad guy is chiding the hero for his anti-robot attitude, and asks 'you think they should have banned the internet to keep the libraries open?'

Never past its use-by date

I stumbled across *Shelflife* this week, while looking for some information about access issues for libraries, archives and museums. It is published weekly by the Research Libraries Group (RLG), and is a newsletter for information professionals, relating to RLG's major initiatives. You can subscribe to the service, or look through its archive all the way back to when it started in May 2001. You will find it at http://www.rlg.org/en/page.php?Page_ID=515.

A different way of looking at it

There is an intriguing story in *The Register* at http://www.theregister.co.uk/2004/07/26/web_adaption_technology/. It is about a project called Web Adaptation Technology (WAT) that IBM has started, to make web pages more accessible to people with disabilities. They realise that most pages will not meet W3C accessibility standards for many years, so they are trying another tack. Through this system, people can make changes to colours, background, text sizes, line spacing and fonts. Users can also choose to eliminate banner advertisements and other images and adjust keystroke timings. Of course, we should continue to make our sites more accessible, but in the meantime this will help many people. The story has a mail link for anyone wanting to contact IBM to participate in the project.

The browser wars, again

There was a time when Microsoft's Internet Explorer (IE) had to struggle against the pack of other browsers, notably Netscape Navigator. That has not been the case for a few years now, and IE became the undisputed leader. But, as the poet Dylan said: the times, they are a-changin'. The security

problems with IE have become so apparent that the US-CERT security group advised in June that people should use a different browser. The most likely competitor to IE is Firefox (<http://www.mozilla.org/products/firefox/>) but, as columnist Brian Livingston advises at http://itmanagement.earthweb.com/columns/executive_tech/article.php/3383081 you may need to make some changes to your sites so that they can be read by Firefox users, and also by those who have other non-Microsoft browsers.

Where shall we put the keywords?

I am always interested in research on how to get the most out of search engines — despite the fact that the search engine developers will then read this research and make adjustments to their systems accordingly. At <http://www.webcredible.co.uk/user-friendly-resources/search-engine-optimisation/keyword-location.shtml> there is a report on an investigation by Jon Ricerca in which he looked at Google, and eventually determined that keywords performed best if they were near the top or the bottom of the page.

UK web archive

The UK Web Archiving Consortium at <http://www.webarchive.org.uk/> has been set up to work on a project for selective archiving of UK websites. Each member of the consortium will be using the NLA's PANDAS software to archive sites relevant to their interests, once appropriate permissions have been obtained from website owners. And there is the rub. If they want to extend the project to archive the whole UK Web, obtaining permissions would take too much time, so the British Library is working with the government to extend the existing copyright law to allow them blanket access to all websites.

Meanwhile, over the pond...

The US Federal Depository Library Program is falling behind in cataloguing and preserving access to government documents published only on the web, and this is a real problem for their depository library program. The Government Printing Office is looking at web harvesting technology to address the problem, but there are concerns that the Deep Web, where most of these 'fugitive documents' can be found (even briefly), cannot be handled by such technology. And a recent California Digital Library study found that about 85 percent of the Deep Web is in the .gov domain.

Getting there

The Royal Geographical Society (<http://www.rgs.org/>), through its 'Unlocking the Archives' project, has opened a new study centre in London with an online search facility. There is also a range of material published online, including factsheets and images from the collections, on top-

Libraries are the perfect environment for robots...

ics such as exploration of Mt Everest and Antarctica.

Beyond usability

We are now seeing the results of all the promotion of usability as a concept — not that everything is perfect, just a little better — so now it is time to look at credibility. There is no point in having the most usable site in the world, if your users do not trust what is on it. The article at <http://www.webdevtips.com/webdevtips/article.php?item=60> explains why credibility is important and describes a number of ways that you can demonstrate that you can be trusted.

On or off, that is the question

A recent study looked at the reading habits of internet users across five document types — journal articles, news, newsletters, literature, and product information. The results are not too surprising, when you consider why people might be reading the different types of material: journal articles were mostly reported to be read in printed form, while others such as online news, newsletters, and product reviews appear to be read mainly online. In particular, users said that they tend not to use online facilities for reading literature. Some of the factors influencing their decisions about whether to print a document or read it online were size, importance, and the intended purpose of document. The full report is at http://psychology.wichita.edu/surl/usabilitynews/62/online_reading_survey.htm.

The death of search engines?

Rita Vine's article in LLRX at <http://www.llrx.com/features/deathsearchengine.htm> considers the current state of the search engine world, and comes up with some provocative conclusions. She says that searching is becoming harder and less productive, and she should know, because she teaches web searching. But I am not sure. I do a lot of searching — at work and at home, professionally and for my personal interests — and I am not finding it noticeably harder to find answers. Of course, with more information online, there are more answers than there used to be, so maybe that is her point. One of her key suggestions is that organisations are becoming more likely to pay for useful material — which is good news for the professional searchers. The LLRX site itself is well worth visiting, providing information workers with news and commentary on a range of internet research and technology-related issues.

Clearing the data smog

The ACRL site at <http://www.ala.org/ala/acrl/acrlissues/acrlinfo/it/informationliteracy.htm> is a good starting point if you are looking for resources and

discussion on information literacy. The Overview section gives a clear and concise introduction to the subject, and then a guide to what you can find on the site, including standards, resources, professional activities and news.

When it is not quite the sincerest form of flattery

The Copyscape service (<http://www.copyscape.com/>) can be used to identify any sites that may have copied your content. It is powered by Google, so only pages indexed by that search engine will be detected, but you can get some interesting results. I tried it on a couple of my pages, and found that one of them had been harvested for an Open Archive Initiative site, which then displayed the metadata elements that I should have put on my site in the first place. But back to Copyscape — it not only tells you where the copies can be found, but offers a 'Whois' link so that you hunt down the perpetrators.

Can you spot the hook?

One of the nastiest tricks on the web is the practice of 'phishing' — fooling people into entering their logon and password details on sites pretending to be a bank or service provider. We do not know the extent of the problem, as the banks are not saying how much they and their customers have lost, but it is a real problem and we all need to be really careful out there. One way to sharpen your wits is to try the Mail Frontier Phishing IQ test at <http://survey.mailfrontier.com/survey/quiztest.html>. It is not as easy as you would think.

For audio-visual cataloguers

Online Audiovisual Catalogers (OLAC) is a group that provides a way to exchange information, to assist in continuing education, and to facilitate communication among cataloguers of audiovisual materials. One of the features of their site is a list of links at <http://ublib.buffalo.edu/libraries/units/cts/olac/websites.html>, including resources from the ALA's Association for Library Collections and Technical Services, and links to other organisations concerned with cataloguing media and metadata, and also to several vendors' sites.

British newspapers online

The British Library has started a project to digitise one million pages from 19th century newspapers. When it is completed, around the end of 2005, users will be able to search through the web and display news, editorials, features, advertisements and photographs. More details are at <http://www.bl.uk/cgi-bin/press.cgi?story=1431>. ■

*There is no point
in having the most
usable site in the
world, if your users
do not trust what is
on it...*