

## New search at Yahoo!



**Kerry Webb**

kerry.webb@alianet.alia.org.au

*High-quality categorisation is crucial if users are to be able to find the right answers in repositories of articles and documents that are expanding at phenomenal rates...*

A few weeks ago I needed to find some lyrics for an old song. No problem: I searched on Google and got several hits, all telling me where I could buy the sheet music. Then I heard about Yahoo's new search facility (they used to use Google technology but now have their own) and tried that. The results provided a very pleasant surprise. The first link in the list was the cover sheet of the digitised version of the sheet music in the National Library's collection, which had links to subsequent pages containing all the words and music. Try the Yahoo! Search at <http://www.search.yahoo.com>.

### What's new in eGovernment

Cheryl Hardy at Multimedia Victoria maintains the eGovernment Resource Centre at <http://www.egov.vic.gov.au>. It is a good collection of links to news and resources from all over the world on matters relating to electronic government services. One of its best features is a weekly 'what's new' alert service, for which registration is free — and easy — at <http://www.egov.vic.gov.au/whatsnew/whatsnew.htm>.

### Taxonomies for better retrieval

In one of the recent editions, Cheryl noted a White Paper from Factiva titled 'Making solid business decisions through intelligent indexing taxonomies'. While its main purpose is to showcase the features of the taxonomy used in Factiva's information services products, the paper does give good examples of how such technology can be generally used in applications of this type. As they say, 'High-quality categorisation is crucial if users are to be able to find the right answers in repositories of articles and documents that are expanding at phenomenal rates'. See the White Paper at <http://tinyurl.com/2fd3y>.

### Musings on searching

Tim Bray is a Canadian who has been involved in software development — especially in searching — for many years. He has put together 'On search, the series', a collection of essays that he wrote last year, and you will find them at <http://tinyurl.com/3cua9>. He writes well and does not go into too much detail, but there is just enough to entertain both the beginner and those who fancy that they know a fair bit about the subject. The series covers subjects like precision and recall, fuzzy matches, metadata (of course) and system interfaces.

### Ee ba gum

The British Library has just put online a collection of northern English accents. This collection shows the huge diversity of accents over the north (and indeed,

even within one county), but many of the recordings also provide a glimpse of social history in these areas. Listeners can hear what someone from Barnoldswick in Lancashire sounds like, or a Hartlepoolian or a man from Man. It is at <http://www.collectbritain.co.uk/collections/dialects/>. Ecky thump!

### Footers

Here's an article by Jeff Lash, a usability specialist and columnist, being a little provocative: <http://tinyurl.com/3x8pu>. The general wisdom about home pages (enshrined in many sets of official guidelines) is that each must all fit on one screen, without any need for the user to scroll down to see everything. Jeff does not agree. He points to the appearance of many new features on popular sites where the space at the bottom of the page is being seen as quite valuable. Rating systems, for example, are being placed there by Yahoo!, and Amazon is now offering special deals advertised in the footer: a sure sign that people are taking the time to go to the bottom of the page.

### OAI Workshop in Geneva

In February, CERN held another in its series of workshops on the Open Archive Initiative (OAI). It brought together librarians and information specialists, publishers, scientists and university managers who were interested in delivering the benefits of open archives technology and open access publishing to libraries. The conference intended to develop priorities for initiatives to be undertaken, in order to increase the impact of OAI. Various papers have been published at <http://agenda.cern.ch/fullAgenda.php?ida=a035925>.

### ...and something similar in Southampton

The CERN meeting was followed a few days later by an international meeting on Open Access Provision for University Research Output. There was a strong contingent of Australians (including those currently living in Scotland — hi Helen!) in attendance and presenting papers. The presentations can be found at <http://opcit.eprints.org/feb19prog.html> and the Declaration of Institutional Commitment from the meeting is at <http://www.eprints.org/signup/sign.php>.

### Time to start preserving our heritage

The National Library has recently published *Preserving Australia's multicultural documentary heritage: a starter kit*. It has been compiled with diverse community groups in mind, both those with an existing collections and those just starting out. It has thirteen fact sheets covering areas such as collection management and preservation, disaster preparedness, and making

the collection accessible. It also describes three existing multicultural archival as case studies. See it at <http://www.nla.gov.au/multicultural/kit/>.

#### **The creative commons**

There are plenty of stories about native tribes who do not like their photos being taken, but what if there was a law in place that said that you need someone's permission before taking their image? In a report in the e-magazine *spiked*, Lawrence Lessig is quoted on what the consequences might have been for photography and film if they had been regulated as closely when they first emerged as the internet is today. His speech was part of the UK launch of the Creative Commons initiative, which aims 'to build a layer of reasonable, flexible copyright in the face of increasingly restrictive default rules'. The report is at <http://tinyurl.com/224vr>.

#### **A meta-announcement**

Tony Barry has announced that announcements are back on the agenda. Until around eighteen months ago, Ozlib Announcement was a prime source for circulating information on events that relate to Australian libraries and librarians. At that time, technical problems caused it to fold, but now it is back on the air. It will continue subjects such as events, appointments, developments, products and conferences. To find out more about it, and to subscribe, see <http://tony-barry.emu.id.au/lists/ozlib/>.

#### **Well, it's been a long time coming**

The Egyptian government and IBM have worked together to create 'Eternal Egypt', a project to provide an extremely rich repository of information about the country's cultural history. With IBM providing the funding and the Egyptian government contributing the experts who produced the content, the site at <http://www.eternalegypt.org> provides an impressive display of Egyptian heritage. After the opening page — with the eerily gliding pyramid — you can pursue a range of navigation mechanisms for the site: timelines, topics, types of artefacts, connections and searching. A nice feature is the extensive use of IBM's text-to-speech system to provide a spoken version of the text on the page.

#### **Problems at the UK e-Learning University**

The BBC has reported at <http://tinyurl.com/2dpk6>, that the UK's international online university is having a little difficulty. The university was established in 2001 by the UK Government, with contributions from several British universities and a number of major technology companies. The main problem appears to be that many of the

universities contributing to UK e-Universities Worldwide (UKeU) have also been offering their own online courses, and it also seems that many of the overseas students who had been expected to study at UkeU over the internet have actually travelled to Britain to enrol in university courses. It is another implication of being able to shop around online.

#### **What users want**

It is funny how long it takes us to appreciate the blindingly obvious. Mark Hurst is another usability expert, and he has developed the page paradigm, which states that users will do either of two things on a web page: follow a link that takes them closer to their goal, or click on the back button. OK — it is a simplification, but still valid. But there's more: he also believes that users don't care where they are on a website. Which makes me wonder about site maps. Are they there just to make use weenies feel better? Like Advanced Search, that hardly anyone ever uses? His article is at <http://tinyurl.com/yr5mb>.

#### **Web services at LC**

The Library of Congress has begun using the OpenURL draft standard to deliver Web services for its *Handbook of Latin American studies online*. OpenURL automatically hooks up online bibliographies and citations to web services such as text databases, search engines, catalogues and other resources; it also standardises the syntax for metadata, unique identifiers and contextual information. While bibliographic citations normally give only the author's last and first names, book title, place and year of publication, and publisher's name, an OpenURL tag at the end of the citation could instantly link the searcher to related web services: viewing the full-text of the book, placing an order for it, looking up the author's other works, seeking online reference help and so on. There's more information at <http://lcweb2.loc.gov/hlas/openurl.sysadmin.html>.

#### **Deep web research**

The Deep Web represents a lot of material — around 600 billion pages of information all over the place in files and formats that the current search engines on the internet either cannot find or have difficulty accessing. While the major search engines continue to probe this untouched area (the story above about the NLA's digital collections is an example) there's still plenty to do. Marcus P Zillman has done a lot of work in cataloguing the research in this field, and you will find a summary at <http://www.11rx.com/features/deepweb.htm>. ■

*...what the consequences might have been for photography and film if they had been regulated as closely when they first emerged as the internet is today...*