

From the conference circuit



Kerry Webb

In June I attended the ACM Digital Libraries '98 conference in Pittsburgh (<http://www.ks.com/dl98/>). I had been to the first of the series in 1996 and had been disappointed that there was a degree of hostility between the computer science research community (who of course are the members of ACM) and those of us who work in libraries. I am happy to say that there is less hostility, but there is still a great deal of work to do in bringing the communities together. One example — following the conference, there was a two-day summit on International Co-operation on Digital Libraries, which ran on the first two days of the ALA conference in Washington DC. (A series of position papers were published before the conference at <http://ks.com/icdl/208.html> and a report from the summit has been promised). Just before the conference there was a meeting of the D-Lib Working Group on Digital Library Metrics, a group which has been communicating by e-mail for some time to try to develop 'a consensus on an appropriate set of metrics to evaluate and compare the effectiveness of digital libraries and component technologies in a distributed environment'. The group also presented a workshop following DL '98. Details on the work of the group, its discussion list and its activities can be found at <http://www.dlib.org/metrics/public/metrics-home.html>

One session at DL '98 featured developers of search engines talking about their systems. While one speaker (who should have been more far-sighted) provided nothing more than a promotional spiel for his product, others were more forthcoming. Marc Krellenstein from Northern Light (<http://www.northernlight.com>) discussed how their product had evolved and how they were concentrating on providing a service for more sophisticated users, as well as giving access (for a fee) to non-Web resources. They all agreed that 'spamming' the search engines was a problem and that they are devoting a lot of effort to resolve this. There was also debate on the best way to implement relevance feedback. One method — asking for 'more like this' usually resulted in a more diverse and much less precise result set.

After Pittsburgh, I travelled to the ALA conference, which was big (more than 20 000 delegates and 740 exhibitors, including a well-attended Australian stand). With so many sessions, it is hard to get an overall feeling for what was happening, but each edition of the daily newspaper and other reports can be found at <http://www.ala.org/events/dc98/live/live.html>.

In her inaugural address, ALA President Ann Symons indicated that ALA would continue to be active in pursuing the freedom to read for library patrons of all ages. Further activity in this area is supported by a recent report by the American Civil Liberties Union, available at <http://www.aclu.org/issues/cyber/box.html>. The report

discusses why blocking technologies are wrong for use in public libraries and also for ISPs. This is quite relevant to the Australian experience, because Senator Alston announced in January that the Federal Government was considering internet content regulation and that a consultancy had been established to advise on technology such as filtering software and blocking devices. He said in a speech in Sydney in May (<http://www.dca.gov.au/speeches/kids.html>) that he expected a report from the ABA by the end of June on the protection of children from unsuitable and harmful material in the online environment. Let's hope they have done their homework.

The UMI vault

At the ALA conference, UMI announced the Digital Vault Initiative, a massive effort to digitise its 5.5 billion page archive containing documents going back to the first texts printed in English. The project will take several years and when it's finished it will be one of the greatest resources of historical documents on the Net. For more information see <http://www.umi.com/hp/Features/DVault/>.

NewHoo

There's a new directory to internet resources that's being built by a staff of volunteer editors. It's called NewHoo, and the idea is that by enlisting people from across the Web to serve as editors, the service will better be able to keep up with new entries and changes that produce backlogs and dead links on other services such as Yahoo. The site now boasts over 400 editors, who have categorised 31 000 sites into 3900 categories — much less than Yahoo's 750 000 sites, but the use of people with expertise has produced some areas where the listings have more depth than at Yahoo.

It could work, but it depends very much on the continued contribution of the editors and their sense of professionalism. See it at <http://www.NewHoo.com/>.

Information Technologies and Disabilities

In the current issue of this journal (at <http://www.rit.edu/~easi/itd.html>) the articles include a review of CD-ROMs for mathematics and science and how useful they are to people with vision impairment, accessible textbooks for K-12 maths and science education, and the use of recorded books along with texts for students with learning disabilities. To get the journal delivered in e-mail, send a mail message to listserv@maelstrom.stjohns.edu with sub ITD-JNL firstname lastname.

Interfaces to scientific data archives

The US National Science Foundation conducted a workshop in March to examine ways in which users could access scientific data using the sophisticated, customised data-mining and knowledge extraction tools which are necessary for the exploitation of these resources. Topics discussed included the use of metadata (especially the use

I am happy to say that there is less hostility, but there is still a great deal of work to do in bringing the communities together.

of the XML language), distributed computing, relational vs object databases, and text vs GUI interfaces. The full report (in various formats) and supporting material can be found at <http://www.cacr.caltech.edu/isda>.

But what is this XML?

It is being suggested that XML (the eXtensible Markup Language) will be the next big thing on the web. It's not a markup language like HTML, but a meta-language that can contain markup languages in the same way as SGML. For an introduction to XML, you can 'X Marks the Spot: eXtensible Markup Language opens the door to a motherlode of automated web applications' at <http://www.cs.caltech.edu/~adam/papers/xml/x-marks-the-spot.html>.

ISP leaving Germany

A couple of months ago, there was a court ruling in the German state of Bavaria which held the former head of CompuServe's German operations responsible for allowing the spread of pornography over the commercial service's network. As a result of this, the US ISP PSINet has now decided to move part of its internet operations out of Germany. PSINet will maintain a strong presence in Germany, but will move a part of its business that stores and displays web pages for home users and clients to other parts of Europe.

Automatic cataloguing and searching contest

The Fritz Kutter-Fonds was established in 1988 to encourage work in information processing and applied computer science. It has announced a special contest with a prize of SFr10 000 for innovative methods in automatic cataloguing. The contest involves a sample set of data based on a set of 500 books of Elias Canetti's private library in the Zentralbibliothek Zürich, where those pages of each book that a librarian considers relevant for cataloguing have been scanned. The images of the scanned pages and OCR-converted versions of each image will be provided. Contestants will be required to answer queries representing a specific information need as well as they can. Data will be available early in August, with results due at the end of March 1999. Details of the contest are at <http://www.kutter-fonds.ethz.ch/contest99.html>.

Kodak PhotoCD technology

Cornell University Library's Department of Preservation and Conservation has just published a brochure summarising the findings of a study to evaluate the use of Kodak PhotoCD technology for preserving and making available a range of research materials. The study was conducted in co-operation with the eleven New York State research libraries, with six other institutions investigating the applicability of the findings. The project was designed to evaluate the PhotoCD technology by controlling the factors affecting image quality during photography, digitisation, and on-screen viewing. The study involved only

paper-based documents, and was limited to documents scanned using the basic Kodak Photo CD method (Image Pac). See the report at <http://www.library.cornell.edu/preservation/kodak/cover.htm>.

New at LC

The Library of Congress National Digital Library Program released four new collections in June. They include documentation of a Nevada cattle-ranching community created by the Paradise Valley Folklife Project with a focus on the family-run Ninety-Six Ranch, a collection of documentary photographs from the Great Depression to World War II, railroad maps from the period 1828–1900, and a collection of more than two hundred social dance manuals. You can find all of these on the American Memory page at <http://memory.loc.gov/ammem/>.

More on Y2K

Sun Microsystems has published a new white paper, 'Are you ready for the Year 2000? Computing impacts of the new millennium on educational institutions'. It provides a succinct analysis and recommendations to campus IT directors and librarians trying to understand and cope with the Year 2000 issue. HTML, PDF, or Postscript versions are available at <http://www.sun.com/edu/admin/whitepaperindex.html>.

All things web

Recent discussion about the ABA Web site (through the use of a proprietary Web authoring software they had created pages where links were not highlighted) pointed to 'All things Web' at <http://www.pantos.org/atw/>. It is a very good site, containing a lot of information about Web usability. I like its introductory quote 'Always and in everything, remember your reader' and later Einstein's 'Things should be made as simple as possible — but no simpler'.

Z39.50 list

The Texas State Library and Archives Commission has created a list for the discussion of Z39.50 implementation and use in the state of Texas, although people from the rest of the world are welcome to join — heck, the state's big enough for a whole passel o' folks. To join, send a message to lstproc@link.tsl.state.tx.us with `subscribe z3950-tx firstname lastname` in the message.

Lexis-Nexis on the move

Patrons at more than 600 United States university libraries will get access to a limited version of Lexis-Nexis in a new deal announced in July. The Academic Universe service contains fewer sources than the parent database but is available via the web and requires no special training. The one catch is that Lexis-Nexis will sell advertising on the service, but rest easy — they've specifically excluded alcohol and tobacco. That's not surprising — I don't recall having seen those products advertised on general Web pages. ■

*It is being suggested
that XML will be
the next big thing on
the Web. It's not a
markup language
like HTML, but a
meta-language that
can contain markup
languages...*

The ALIA home pages and the ALIA web service, ALIANet, can be found at the following URL:
<http://www.alia.org.au/>

**SAVE YOURSELF
THE TROUBLE OF
TYPING THESE
URLS!**

This column (with associated hyperlinks) is on ALIANet, and has the URL of <http://www.alia.org.au/incite/www/>

[Please note that full-stops at the end of any URL or e-mail addresses, or underlines, appearing in *inCite* should NOT be used]

Contributions and suggestions for this column are always welcome. Please e-mail Kerry Webb at kerry.webb@alianet.alia.org.au or via the ALIA National Office, incite@alia.org.au