eLucidate

Vol. 2 Issue 2, March/April 2005

ISSN: 1742-5921

Contents

Report on the use of an award from the John Campbell Trust to attend the Annual Meeting of the American Society for Information Science & Technology

I applied for a grant from the John Campbell Trust in January 2004, in order to help support participation in the Annual meeting of the American Society for information Science and Technology, to be held in Providence, Rhode Island in November 2004. My intention was to make an intellectual contribution to the meeting, presenting a paper, *A labor theoretic approach to information retrieval*, and to participate in society governance, including acting as Joint International Liaison Officer.

I was able to attend the meeting and to participate fully.

The meeting confirmed the intellectual diversity of the society. Presentations revealed broad, and varying, conceptions of information science. Colleagues known through their publications became known by personal contact, and, in turn, personal dialogue may stimulate further presentations and publication.

A number of interesting, although necessarily entirely novel, intellectual issues were raised. A keynote speaker, J.C. Herz, questioned the assumption and method of post-Enlightenment thinking, of breaking issues down into their smallest parts for analysis, was still tenable. Keeping things found was not as well studied as retrieving information (reminiscent of David Lodge's observation on literature being romance and meetings and life about marriage). Mechanisms for forgetting were also alluded to, although Freud's central question – why should anything be forgotten? – was not discussed. Other sessions noted that the provisions of the Patriot Act had long been on the political agenda. Concepts of group size in relation to the limits of human cognition and the connections between open source information and the public sphere were also considered.

One of the most vital and vitalizing sessions was the Global Information Village Plaza. Issues for consideration, many posed as What if? questions, had been circulated in advance and responses were discussed in the session. What if journal submissions were evaluated for their contribution to knowledge rather than their formal correctness? What was the impact of the information society, at both a professional and personal level? Responses revealed a low level of immediate awareness of the pervasiveness of ICTs as control mechanisms.

Society governance and the conference sessions were partly intermingled (business meetings are also open to all society members, apart from deliberately closed sessions). One presentation gave an account of a survey of ASIS&T members. Their median age had increased slightly since the last survey in 1979 (with reference to the membership not to the surviving individual members, although tautologically true, if understated, of them as well). There was a need for younger members of ASIS&T. Information technologists remained uninterested in ASIS&T. Library and information studies was dominant as members' highest degree.

Formal governance activities also occurred. Some meetings were concerned with issues relevant to European and international members. There was scope for the growth of the European Chapter, with a broad understanding of Europe, and for further internationalization. Committee charges for the 2004-2005 year were reviewed. The growth and size of the information science field was revealed by the increase in the number of papers, already substantial, submitted to JASIS&T. There was a growing professionalisation of information architecture, as a sub-discipline of information science, although also with other connections. A reciprocal sponsorship agreement had been established with Communication and Information Technology Section of the American Sociological Association.

International themes formed a persistent focus of attention. The meeting was marked by the number and range of international visitors. International issues were attracting increasing interest and sympathy among ASIS&T members. The International Reception was well attended. International members were valuable to the society and increasing internationalization of the society was envisaged. International standards represented one significant area where there was need for involvement.

Some innovations to the format of meetings had occurred in recent years. A poster session had been introduced and was expanded for the 2004 meeting. Posters were well attended, with a number of lively and informative dialogues between presenters and other attendees.

The next year's meeting will be Charlotte, North Carolina in November 2005. The conference proceedings will be published in electronic, rather than print on paper form, and will be searchable and include panel presentations. The Global International Village Plaza should be developed further and remain interesting.

I would like to record my appreciation of the John Campbell Trust in assisting me in participating in a successful meeting.

Julian Warner
School of Management and Economics
The Queen's University of Belfast
j.warner@qub.ac.uk

The New Wave of Bookmarks

February 2005 Search Tip of the Month from Mary Ellen Bates

Every good researcher maintains at least a small collection of bookmarks of her favorite web resources. It's the virtual equivalent of the "ready reference" shelf of books that librarians use to answer the majority of their questions. And some researchers (and you know who you are) have bookmark files that run into the hundreds. While it is nice to have handy access to key resources, it can be difficult to manage and, more importantly, find information buried in a large bookmark file.

A new solution to this problem is Furl (http://www.furl.net), a "social bookmarking" service. What is social bookmarking, you ask. The idea is similar to the social networking services such as *LinkedIn* or *Orkut*, or *Amazon.com*'s "people who bought this book also bought these books." If other people have bookmarked the same site you just bookmarked, perhaps you would be interested in seeing the other sites they bookmarked as well.

I am not convinced that this is an efficient resource discovery technique, but the features of Furl have certainly caught my attention. In a nutshell, Furl lets you bookmark pages to the Furl site, annotate the bookmark, and assign keywords ("controlled vocabulary" to us librarians) to the bookmark. Furl not only stores that information in your My Archive area of its site, but it also archives a copy of the page as it exists when you bookmark the page. Whoa! That means you can, in essence, build your own mini version of the Wayback Machine (http://www.archive.org). While the personalization features of MyYahoo, MyJeeves and other search engines offer some annotated-bookmark features, these are generally tied to marking pages found through a web search results page, and none of the my-[search engine] archives the full text of the page being bookmarked.

What else can you do with your "Furled" pages? You can run a search of your bookmarks, which actually goes out and searches the current version of all the bookmarked pages, not just your annotations of that page. Work with me here... what that means is that you can build your own searchable version of the web, or at least that corner of it that you are particularly interested in. (If the search terms don't appear in any of your Furled pages, Furl returns the first 300 search results from a *LookSmart* web search.)

You can also choose to search the universe of Furled pages – not just your own but those of all other Furl members. The search results page shows the links to Furled pages and a snippet showing your search term in context, the number of other Furl members who have Furled the page (an interesting indication of the relative influence of that page, at least within the Furl community), and a link that lets you add that page to your own Furled bookmarks. One more click and you can see comments people made when ~they~ furled the page; again, a potentially useful perspective into what other Furl members are thinking. Interestingly, the results of an entire-Furl search are relevance-ranked based on PeopleRank (note the similarity in naming to *Google*'s PageRank). The more Furl members that Furled the page, the higher its PeopleRank.

So, what do I use Furl for?

* I see a thought-provoking article and think, "hmmmm, looks interesting, but I'm not going to have time to read it in the next few weeks. It's likely to disappear from this newspaper's site before then. Rather than print it off, I'll just Furl it and read the archived copy later."

* I am conducting web research and want to save a large collection of white papers, reports, or other resources. While I can save them to my own computer, I want a copy on another server that I know is backed up daily.

* I see a web site that I know I will need to use in a month or two. My short-term memory isn't what it used to be; I would rather Furl it, with a long note reminding myself of why I care about this, than try to find it a month from now.

* I am preparing for a presentation that includes demonstrations of web sites. While I can build an HTML page of the URLs I want to show, loading the bookmarks on Furl ensures that, even if the site I want to show is down, I can show the archived page from my Furl collection. I can even filter my Furled bookmarks by topic; for a recent presentation, I loaded all my bookmarks into Furl and tagged them with a topic name of "Melbourne 2005". When I started my talk, I just filtered the bookmarks for that topic and the attendees never had to see my other Furled bookmarks.

* I want to share my personal view of the web with others. While I could always set up a blog, a different approach is to let my friends and colleagues subscribe to an RSS feed of the Furled bookmarks (and my associated annotations) I chose to make publicly available.

I can think of a number of other uses for Furl, but I encourage you to head over there and try it out yourself. One suggestion, though: while you can set the default so that your Furled bookmarks are marked private, play nice and publicly bookmark the resources you value and that do not violate confidentiality (that is, that don't disclose anything to your organization's competitors if they see you have bookmarked it). For example, if your company is considering entering a new market, it would be best not to disclose the bookmarks pertaining to that market. For other

bookmarks, though, share them with others. The continuing value of Furl is not only in its personal web features but in the shared knowledgebase and recommendation system of the social bookmarking network.

Note that another social bookmarking service that has gotten some good reviews lately is *del.icio.us* (yes, that's the URL: http://del.icio.us). It offers some nice features if you are interested in seeing what other people have bookmarked, including an RSS feed of bookmarks added to del.icio.us on a specific topic. However, as Joshua Schachter, the creator of del.icio.us concedes, "This system is pre-pre-alpha; many features have yet to be added. Additionally, many, many bugs remain. Please be careful." (That sounds like those signs you see in gift stores: you break it, you buy it!)

My primary objection to del.icio.us is that there is no straightforward way to keep your bookmarks private. The strong default is to share your bookmarks and their associated "tags" with other subscribers. While I understand that this is the way that a social bookmarking system grows, to virtually require that all bookmarks be public reduces the system's usefulness for many researchers. If Schachter adds the ability to set the user default to private, rather than public, bookmarks, and if *Del.icio.us* gets past pre-pre-alpha, it would be a real competitor to Furl.

An RSS feed for my Search Tip of the Month is at http://www.batesinfo.com/tip.rss

What's New in Search Tools or "Alas Poor Google..."

This article is based on presentations given by Karen Blakeman at the ISG Meeting "To Google or not to Google" on 23rd November 2004, Cambridge UK and the Information Masterclass at Online 2004, Tuesday 30th November.

This review of what has been happening over the last year looks at Google's new features, the alternatives to Google (and yes there are plenty!), and desktop search (a really hot topic at present).

What has Google been up to? http://www.google.com/

Apart from its stock exchange listing, Google seemed to go into hibernation during the summer of 2004. There was little in the way of new features and it looked as though its database was not being significantly updated. Then, in response to the launch of Microsoft's new beta search engine, Google doubled its database to 8 billion pages and a new numeric range search was added to the Advanced Search page. It did not stop there, though. Google then launched Google Print, Google Scholar, Google

Libraries, Google Desktop Search and Google Suggests to name but a few.

Google Print http://print.google.com/

Google Print is aimed at publishers both large and small. Books are supplied by authors and publishers as hard copy and are scanned in by Google. Type in a search in the format 'books about.....' and Google lists those that match at the top of the results list. It appears that there is a maximum of three books per search but that could just reflect the number of books that have been supplied by publishers. Google displays an image of the page in the book that mentions your terms the most, and you can view a maximum of two pages either side. There are also links to bibliographic information, to book stores or the publishers own site where you can buy the book.

The Google Print is programme is now being expanded to include selected books from the libraries of Harvard, Stanford, the University of Michigan, and the University of Oxford as well as The New York Public Library. Google Print only works on the Google.com site and not on any of the country versions.

Google Scholar

Google Scholar (http://scholar.google.com/) enables you to search for "scholarly literature including peer-reviewed papers, theses, books, pre-prints, abstracts and technical reports from all broad areas of research." A wide range of academic publishers are covered such as professional societies, pre-print repositories and universities, as well as scholarly articles available across the web. There is no source list so it is down to guesswork and experience in working out what is covered.

The advanced search options are limited to author, journal and date of publication. The author search is a bit hit and miss and varies depending on the format in which you type the name. 'K H Blakeman' and 'KH Blakeman' (without the space between the K and H) give different results.

Google Scholar automatically analyses and extracts citations. This means that your results may include citations of older works, books or other off-line publications. The results are listed by relevance, which is a disappointment for those of accustomed to sorting this type of literature search by date, author etc. Google Scholar is still in beta so there is hope that at least a date sort will be added.

Google Suggests

Google Suggests can be found on the Google Labs page (http://labs.google.com/). Simply start typing in your search and Google comes up with a list of suggestions for completing your strategy together with the number of results each will give. My search on gin

and tonic came up with 'gin and tonic recipe', 'gin and the devil', and 'gin and tonic movie'! (See figure 1)

Figure 1: Google Suggests

Limit on number of search terms increased

Google has at long last increased the limit on the number of search terms from 10 to 32. As many commentators have said, it is long overdue and makes it a lot easier to carry out more sophisticated searches and searches for quotations.

Alternatives to Google

So what is wrong with Google? For many searches - nothing. Google comes up with the goods most of the time for the majority of users but there are times when one just cannot find relevant information. This can be because Google, for a variety of reasons, has not covered the pages that are relevant but more often the most relevant information is buried way down near the bottom of your list of results. Different search tools rank the results of the same strategy differently and what is number 22,736 in Google's list of pages may be number 5 in Yahoo's.

Try comparing the performance of different search engines in Thumbshots Ranking (http://ranking.thumbshots.com/). This covers Google, Yahoo, AltaVista, MSN, AltheWeb, Teoma and Wisenut, and you can compare two at a time. Type your search strategy into one of the search boxes and then select the tools you want to compare. Thumbshots compares the first 100 pages found by the search engines and represents them as two lines of circles. The pages that are common to both search engines are represented as solid blue circles and linked by a blue line. (see figure 2)

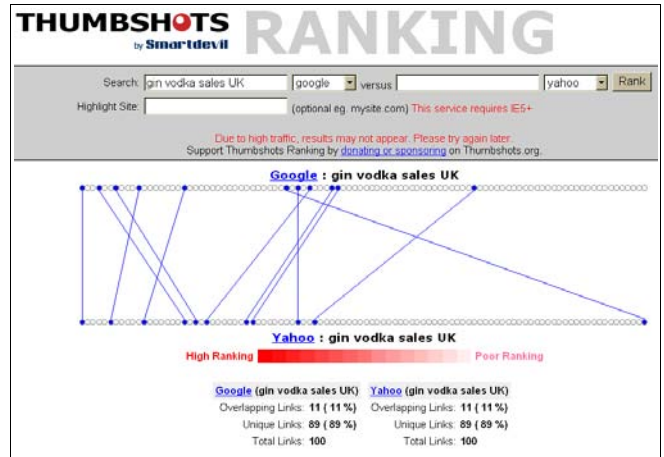Unfortunately, Thumbshots Ranking only works in Internet Explorer.

Figure 2: Thumbshots Ranking

Yahoo http://www.yahoo.com/ or http://www.yahoo.co.uk/

Over the last 18 months Yahoo has bought up Inktomi, Overture, AlltheWeb and AltaVista. At the start of 2004 it dumped Google as its partner for web search and launched its own web database. Yahoo has not revealed the size of its database but it is certainly giving Google a run for its money and it does sometimes gives better results than Google.

How do Google and Yahoo compare?

1. The advanced search features of Yahoo and Google are very similar.

2. Yahoo has the advantage that it supports all three Boolean operators (AND, OR, NOT) and parentheses for nested searches. Google only supports OR and you cannot carry out sophisticated nested Boolean searches.

3. Google has a link: command that finds pages that link to your specified URL, but it will only find links that match your URL exactly and shows only a small selection of linking pages. Yahoo has two commands. 'Link:' finds pages that links to a specific URL, for example *link:http://www.rba.co.uk/search/* (you must include the http://), whilst 'linkdomain:' finds pages containing links to any page on the specified domain, for example *linkdomain:rba.co.uk* (do not include the http:// with this command). The link command is a useful way of finding similar types of pages, the assumption being that pages that link to another are often similar in content.

4. Google indexes the first 100K of a page: Yahoo indexes up to 500K. This may not sound significant when searching plain old html web pages but can make a difference when looking for file formats such as PDFs, which can be very large. Some

search gurus have spotted that Google sometimes drops the 100K limit but it varies from day to day and results are not reproducible. This could be something that Google is working on.

5. As well as the usual PDF and Microsoft Office file formats Yahoo allows you to restrict your search to RSS/XML formats. Perfect if you are searching for news feeds on a particular subject.

6. Google has a number range search.: Yahoo does not. Specify two numbers separated by two full stops with no spaces, and include a unit of measure or some other indicator of what the number range represents. For example DVD player £50..150, or toblerone 1..5 kg. Number ranges can be years, weights, prices, temperature, concentrations etc.

7. Google has a synonym search: Yahoo does not. Precede your term with a tilde (~) and Google will look for similar or related terms. There is no information on how the synonym lists have been compiled and it is not always accurate. This may not matter for more general searches but can be an issue with scientific or technical searches. For example a synonym search on the artificial sweetener ~aspartame will pick up products that contain aspartame and sucralose, which has a different structure and properties.

8. Both Yahoo and Google have an excellent current news search with similar advanced search features, including options to limit your search to one source or to sources from one country. I have found that for most of my own searches Google sometimes has better International coverage, but it does vary depending on the search. Both offer news alerts by email but Yahoo can also provide them as RSS feeds (available only in Yahoo.com at present).

MSN http://www.msn.com/

Microsoft pulled out all the stops as far as publicity was concerned when it launched this but I was very disappointed with it. It has a simple uncluttered default search screen - the current fashion - but I found that results for my test searches were poor when compared with Google and Yahoo. Advanced search features are under "Build a search" and are restricted to the link command, country/region, language, and site/domain. There is no file format command. The News search offers RSS feeds for alerts, which are provided by Moreover.

Exalead http://www.exalead.com/

Exalead was launched in October 2004. With 1 billion Web pages in its index it may seem small in comparison with the the likes of Google, who claim to have 8 billion pages, but the quality of Exalead is excellent and there are some cracking good search features.

For starters it supports wild cards (an asterisk) that can be used to represent any number of letters after a specified string of letters. Then there is word stemming that can be set as the default under preferences, and which finds the variants of a word such as plural/singular and verb conjugations.

If you are not sure how to spell a word try phonetic search or the approximate spelling search on the Advanced Search screen. Another approach is to use the pattern matching feature. Begin and end the "pattern" with a forward slash (/) and use a full stop or period to represent single characters in the pattern. If you are not sure how many characters are in the middle of a word use the full stop followed by an asterisk (*). The help file suggests that pattern matching is ideal for solving [shouldn't that be 'cheating at'?] crossword puzzles!

Full Boolean search is supported (AND, OR, NOT) and there is a proximity command (NEAR) that will search for words within 16 terms of one another. Filetype, country and language search options are supported. RSS and blog search options are promised for the near future.

The results page is pretty nifty too. The usual list of results is augmented with a snapshot of each web page (you can switch this off if you prefer or just display the snapshots), suggestions for related terms, options to view the results by geographical location and file type (for example .doc, .pdf, .xls).

Find.com "True Business Search" http://www.find.com

Find.com was launched in June 2004 and concentrates on business search. It combines results from its own index of business web sites with those from some of the major search engines. It also finds priced reports - "premium research content" - from sources such as Thomson Gale, The Gallup Organization, Frost & Sullivan, BNET, Marketresearch.com, Datamonitor etc.

The Web option is a meta search tool and appears to cover at least Google, Yahoo, MSN, AltaVista, AlltheWeb, SmealSearch, Scrius and About. This option also includes Business Web which are business web sites identified and indexed by Find.com.

Find.com automatically searches for all of your words but supports the standard Boolean AND, OR, and NOT operators all of which must be in capital letters. The standard double quotes can be used to specify phrases.

The Advanced Search screen has options for searching the full text (the default) or titles only, and keyword search (default) or concept. The concept search takes your terms and looks for similar or related

terms. For example, type in automotive and Find.com will pick up references to automobile, motor trade, motoring etc. The drawback is that if you have more than one term in your query you cannot specify which of them is to be treated as a concept. It is all or nothing.

There are also "Threshold" options that enable you to increase the number of documents that are shown from the default of 100 and to change the "minimum relevance percentage" of 10. Worryingly, nowhere on the home page or on the results page does Find.com explicitly tell you that by default it only shows you a small selection of records. This really ought to be made clearer on the home page.

There are several more advanced features but these only apply to the Business Web and Premium Research collections. You can use a question mark (?) inside a word to represent a single character, or an asterisk (*) inside or at the end of a word to represent one or more characters. If you are not sure of spelling use the "fuzzy search" feature by adding a tilde (~) at the end of a single word.

There is also a proximity option. Place your terms inside double quotes and then a tilde at the end of phrase together with a number. For example: "research development"~5 looks for the terms research and development within five words of each other.

You can boost the relevance of a term or phrase in your search by using the caret, "^", symbol together with a boost factor (a number) at the end of the term. For example: "production statistics"^4

Results are sorted by relevance. On the left hand side of the screen is an "Organize results by" column. This enables you to view your results by topic, document format (htm, pdf, ppt etc), site and source. The Topic option clusters results into subject folders and can help narrow down your search.

Answers.com http://www.answers.com/

GuruNet launched its free reference service Answers.com at the beginning of 2005 offering "the best definitions and explanations for over 1 million topics". Information on people, places, words and names is drawn from dictionaries, encyclopaedias and selected web sites. A search on Winnie the Pooh, one of my regular test searches (yes, sad isn't it?) came up with information gleaned from Who2, the New Dictionary of Cultural Literacy, and Wikipedia. For the artificial sweetener aspartame Answers.com came up with definitions from a standard and a medical dictionary, an entry from Wordnet, the entry in Wikipedia including its Chemical Abstracts Number and structure, and translations into various languages. There are links to other topics that mention your term or phrase and to searches on Google for web pages,

images and news. The blogs search option uses Technorati.

If you find you are becoming hooked on the service, as I totally am, you can download the 1-Click Answers bar. Once installed there is an additional feature that automatically runs an Answers.com search when you Alt-Click on any word or phrase. This works on any document in any program on your screen.

Turboscout http://www.turboscout.com/

This is a really neat tool that allows you to search the results of 23 search engines by typing in your search, and then clicking one by one on the engines that interest you. This may sound tedious and I could not at first see the point of it. The really useful aspect of it is that you decide which tool to search in. It could be two or three under the Web tab for example Google, Yahoo, Vivisimo and then you could switch to the Reference tab and look in Wikipedia, FindArticles or Scirus for example. See figure 3.

Figure 3: Turboscout

Desktop Search - the latest hot topic

When Microsoft announced that it was working on a tool that would combine Internet with desktop search, the competition started beavering away to beat them to it. The idea behind desktop search is that as well as, or instead of, scouring the web for information on a topic you can also search documents stored locally on your computer. These "documents" could be cached web pages, email messages, Microsoft Office documents, PDFs. In theory anything. In practice, if you use anything other non-Microsoft products for example Lotus Notes, Open Office, Pegasus Mail, Thunderbird etc. you may find you are left out in the cold.

Blinkx http://www.blinkx.com/

One of the first of the current batch of desktop search launches, Blinkx differs from the others in that it claims to eliminate the need for keywords or complex search methods and uses "self-learning algorithms to understand the context" of what you are reading on your computer screen.

When you first run Blinkx, you can tell it which directories you want it to index. The indexing can take anywhere between a few minutes to a few hours

depending on the number of files on your system. It indexes .doc, ppt, Excel, PDF, txt, html and various media files. As for email, it only supports Outlook, Outlook Express and Eudora. I had Blinkx running for 6 days. At the start, the local documents it found for me were relevant but the web results were useless. The company does emphasise that this is a "concept" as opposed to a "keyword" search tool, but there is no explanation as to how the concepts are identified and analysed. I assumed that Blinkx needed to learn more about what I regarded as relevant, but as the days passed the Internet results were no better and my local hard disk results became increasingly bizarre. I suspect that Blinkx works best with narrowly defined subjects. If, like me, you research anything and everything it loses the plot completely.

Copernic Desktop http://www.copernic.com/

Copernic is well known for its search agents that cover multiple search tools and "invisible web" resources, as well as for its web page tracking program. Installation is straightforward but the first time it starts up it can take a while to index the files on your hard disk, as the case with all of these types of tools. It automatically stops indexing when it detects significant activity but you can pause and restart it manually. By default it indexes the My Documents folder but you can customise it by adding other directories or even whole drives. Copernic Desktop can search PDF, XLS, PPT, DOC, RTF, TXT, WP and HTML files. It also indexes your search history, Internet Explorer favorites, Outlook email and contact lists. The most recent version has added Mozilla and Firefox support. A really neat feature is that under the Advanced Options you can tell it to index other file types and add file extensions that are to be indexed as text files.

Google Desktop Search http://desktop.google.com/

The beta version of Google Desktop Search (GDS) gave the impression that it was launched before it was ready in an attempt to beat Microsoft and Yahoo. It was limited to Microsoft Office formats, Outlook, Outlook Express, AOL Instant Messenger, TXT and HTML but did not index PDFs. The most recent official version has added support for PDF files, Mozilla, Firefox, Netscape and Thunderbird. There is also a plugin that indexes Open Office and Star Office file formats.

Some of the original security and confidentiality issues have been addressed but you need to be aware of their significance and know how to switch various options off. GDS can index cache secure https pages that you view, for example online bank statements: untick the relevant box under Preferences. Password protected and encrypted files can also be indexed: again untick the box under Preferences.

Another problem is that documents remain in the GDS cache on your PC even after you have deleted the original files. So it is possible for anyone using your computer to unearth those embarrassing emails and

online chats that you thought were long gone! More importantly for corporate users, though, is that this feature could conflict with document retention and management policies. There are instructions for removing documents from the cache at http://desktop.google.com/features.html#remove but it is far from straightforward and you have to know exactly which files and emails you want to delete.

MSN Desktop Search http://www.msn.com/

MSN Desktop Search is part of the MSN Toolbar Suite Beta and installs as a browser toolbar and optionally in the Windows taskbar. MSN Desktop indexes Office documents, Outlook and Outlook Express email, photos, music and email attachments. You can specify exactly what you want it to index: email and My Documents, email and all of your hard disc, or individual directories. It spots when you are active on your PC and suspends indexing quite quickly, but there is also a "snooze" button you can use to stop it immediately. The main drawback with MSN Desktop is the lack of support for non Microsoft email clients, browser and document formats.

Ask Jeeves Desktop http://sp.ask.com/docs/desktop/

Yet another search tool that assumes that everyone is using Microsoft Office, Outlook and Internet Explorer. Other files that it currently supports include images, MP3 music, video files. It also searches titles and metadata from Adobe PDF files but not the text. It does not index inside ZIP archive, Outlook attachments, contacts and one odd omission is that it does not index the content of Excel files. This a beta release, though, so additional functionality may be added later. It does take a while for it to realise when you want to use your computer and activity has increased, and there is no quick way of forcing it to cease and desist apart from shutting it down altogether.

Yahoo Desktop http://desktop.yahoo.com/

Yahoo is the latest search engine to launch a desktop search program and, in my view, it is by far the best so far. It searches the usual "popular" file formats such as PDF, Microsoft Office, html etc "plus over 200 more". A full list can be found at http://desktop.yahoo.com/filetypes. It retrieves any file on your computer and displays results as you type. The preview option is excellent and most files look exactly as they should in terms of layout. In my tests, it found all the files containing the text I specified, including files without extensions and my Open Office files. My only gripe is that when it is indexing your files, it takes a while to pause when you start using other programs.
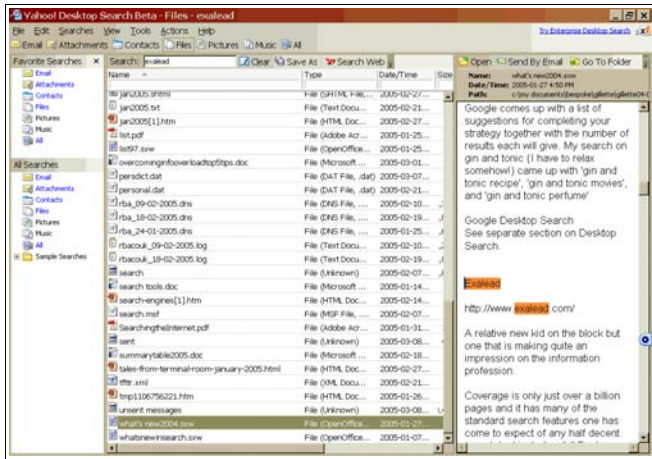
Figure 4: Yahoo Desktop

What next?

Who knows! Google will continue to be the search engine that many people go to first but Yahoo is a offers strong competition. Both continue to add features but it is a question of how relevant these are to us. For example the most recent Google launch was a weather search, but Google decided that only people in the US were interested in conditions outside! New tools will continue to be launched. Many of these will disappear without trace. Others will fail to live up to expectations, and features that promise to be "coming soon" will never materialise.

If I had to pick one to watch, it would have to be *Exalead*. It is a comparative lightweight in terms of database size but has some seriously good search options. If it continues to develop new features and launches its promised Desktop Search it could be a serious contender for Google's crown.

Karen Blakeman
RBA Information Services
karen.blakeman@rba.co.uk

Short Notes

Study shows online citations don't age well

A study conducted by two academics at Iowa State University has shown a remarkably high rate of "decay" for online citations.

Michael Bugeja, professor of journalism and communication, and Daniela Dimitrova, assistant professor of communication, looked at five prestigious communication-studies journals from 2000 to 2003 and found 1,126 footnotes that cite online resources. Of those, 373 did not work at all, a decay rate of 33 percent; of those that worked, only 424 took users to information relevant to the citation. In one of the journals in the study, 167 of 265 citations did not work.

Bugeja compared the current situation to that of

Shakespearean plays in the early days of printing, when many copies of plays were fraught with errors due to the instability of the printing medium.

Anthony T. Grafton, a professor of history at Princeton University and author of a book on footnotes, agreed that citation decay is a real and growing problem, describing the situation as "a world in which documentation and verification melt into air."

Abstract taken from Edupage 14 March 2005; the article is at Chronicle of Higher Education, 14 March 2005 (sub. req'd). See also Oxford Technology Book Predicts Growing Citation Problem

ESRC Society Today

The UK's largest funding agency for research and postgraduate training on social and economic issues - the Economic and Social Research Council (ESRC) - is soon to launch a major new website in the UK.

ESRC Society Today will offer academics, students and researchers unrivalled, free access to high quality social and economic research available, planned and in progress.

As well as bringing together all ESRC-funded research, the website will allow users to cross search material from other key online resources such as SOSIG, the UK Data Archive, the Office of National Statistics and Europa among others.

It will include early findings, full-texts and original datasets and allow users to establish online discussion fora, track down researchers in particular subject fields or find details on ESRC funding and training opportunities. The emphasis will be on accessibility and ease of use and visitors will be able to register for regular news bulletins and email alerts on their areas of interest.

ESRC Society Today - http://www.esrcsocietytoday.ac.uk - is scheduled to go live for a test period in April 2005 before being publicly launched in Summer 2005, replacing the current http://www.esrc.ac.uk and http://www.regard.ac.uk sites.

Librarians are invited to register for further information by emailing societytoday@esrc.ac.uk

Farewell

This is my final issue as Editor of *eLucidate*. Gary Horrocks (gary.horrocks@kcl.ac.uk) will be looking after the Journal whilst the Committee seeks to appoint a new Editor.

Peter Chapman

Online

**Column Editor: Anita Vautier,
Information Services, University of Wales,
Aberystwyth
e-mail: aiv@aber.ac.uk**

Institute of Physics

The Institute of Physics have digitised their journal collection back to 1874. This means that over 110, 000 articles are available digitally as PDF documents that can be cross searched and contain linked references. The archive is available as a one off purchase or on a subscription basis. http://journals.iop.org/

JSTOR

A new search engine has been released for JSTOR. This new search engine has been developed to improve the speed ability. The basic search will now search across all the content which includes articles and reviews. Other features available in the basic search are Boolean operators, phrase searching and field searching. When a Boolean operator is not entered the automatic default is AND. http://www.jstor.org/help/search.html

Ulrich's Resources Linker

Bowker have recently announced their Resource Linker for serials in the library. There is an A-Z listing of journal titles and a search facility that searches the journal title and articles. It is also possible to subject browse. The Resource linker is hosted by Bowker, who will set-up, maintain and update the link server, which will update any catalogue changes made to the collection. http://www.resourcelinker.com

Google Scholar

This search engine is provided by Google and is the beta version. A search entered will retrieve books and abstracts for any articles found. If an open access journal appears in the results the full text will be available. The Google Scholar search engine searches scholary journals and any peer-reviewed literature. http://www.scholar.google.com

People's Network Online Enquiry Service

This service will be available to the public in March. It provides a free information service twenty four hours a day, seven days a week. The public can submit a question via the online form. The service is provided by MLA (Museums, Library and Archives Council) who are currently piloting the scheme. They are asking library and information staff to submit a question to help test the service. http://www.peoplesnetwork.gov.uk/

Internet News

**Column Editor: Susan Miles;
e-mail: susan_miles_2002@hotmail.com**

Googlizing the Library?

Google announced in December 2004 that "it is working with the libraries of Harvard, Stanford, the University of Michigan, and the University of Oxford as well as The New York Public Library to digitally scan books from their collections so that users worldwide can search them in Google" [1]. This precipitated newspaper articles both here and in the U.S. trumpeting the liberation of students and researchers from "musty library catalogues", along the lines of the Guardian's "endless journeys to the library could become a thing of the past" [2].

The Scope of the Project
So, what exactly is the scope of this project? At first glance one's initial reaction is that this new project is certainly a Herculean task. Yet when I tried to find out, actual numbers seem somewhat thin on the ground – whilst the University of Michigan is contributing seven million volumes and the Bodleian Library one million volumes [3], other participants are more guarded about the scope of their contributions. Stanford University has a collection of over eight million volumes but its FAQ on the new project says "The agreement with Google is open-ended; it neither targets specific collections nor specifies a minimum or maximum number of books to be digitised". [4]

Similarly, the New York Public Library's press release makes no mention of how many books will be scanned, it is participating on the basis of a "pilot program that will make a subset of its books available" [5]. There is another pilot project at Harvard University; initially 40,000 of its 15 million volumes will be digitised. Interestingly their selection critieria for materials to be included in the project is to randomly select entire shelves of materials at the Harvard Depository, which will provide a diverse range of items in terms of age, topic and language mainly because the books are shelved by size! [6]

How will it be done?
Google seems to be taking a two pronged approach to the digitising process, Stanford University will send books direct to Google's headquarters in California, whilst the other four libraries will digitise in-house. Google will establish local units which will carry out the digital scanning and processing using proprietary, high-speed scanning stations, which once fully operational will be able to produce as many as 10,000 electronic books per week. Once the book has been scanned, the resulting images will be sent to Google's central facility for quality control and conversion to text. The contributing library will receive a copy of the digital file for their own use. Whilst Google appears to be

providing the staffing for the scanning stations I wonder how much library staff time and effort will be taken up with delivering, keeping track of, and returning, such large quantities of books on an ongoing basis. Other observers have conjectured that an average scan rate of 3,200 volumes a day for 365 days a year will be needed to meet the 6-year timeframe for the University of Michigan's seven million volumes. As far as guessing how much this is going to cost Google, Gordon Macomber, president and CEO of Thomson Gale has indicated that $10 per book was below Gale's experienced cost. They have extensive experience with digitizing for their 18th and 19th Century Collections products. [7]

As anybody who has had anything to do with the digital capture of documents will tell you, these projects can (perhaps this should read often) prove to be more difficult than is initially anticipated, issues such as quality control, or costs, can easily escalate, and let's not forget the challenge of managing the workflow at the sort of work rates which will be needed.

Copyright Issues
Initially it is likely that out of copyright works will be the first books to be digitised. Users searching with Google will see links in their search results page when there are books relevant to their query. Clicking on a title will deliver a Google Print page where users can browse the full text of public domain works and bibliographic data for copyrighted works. It is intended that searchers will read these books online, and may not be able to print out the whole document. There may be plans to offer searchers tasters for copyrighted works, namely, three "snippets" of text from throughout the volume. A "snippet" would consist of approximately three lines of content. There would be pointers to direct searchers to where they can obtain full copies of copyrighted works. By the end of the project Google will have effectively created the largest out-of-print bookstore.

Industry Response
One day after the Google announcement, the Internet Archive announced details of their own collaboration with libraries of five countries [Canada, China, Egypt, India and the Netherlands] to add their digitised books to open-access archives. Over one million books have been committed, with 27,000 currently available, with an additional 50,000 expected in the first quarter of 2005. [8] It remains to be seen if, and how, companies like Amazon, Microsoft or Yahoo! will respond to Google's announcement. Equally, the vast collections in the Library of Congress and the British Library remain outside of these large-scale digitisation projects. They remain the jewels in the crown.

Unanswered questions
What about duplicates? There appears to be no plans at present to prevent duplicate copies across contributing libraries being submitted to Google. Also, it will not be possible to select items by contributing

library as no separate identifier will be created. Since the whole book is being digitised, if a library has barcoded and added its own identifying marks to say, the front or rear pages, then these will be captured and will be readable.

What about the impact on the wider library community? Carol Brey-Casino, current President of the ALA said in a Wall Street Journal interview, "We had this conversation when Internet began to get popular, and what's happened is that library visits have doubled in the last decade". Many commentators see no threat but a further push for librarians to become more active in empowering users to find, evaluate, use and access information in a variety of formats.

Some final thoughts:
I wonder if information professionals will be recruited into these digitising projects, or will existing experts in this field be clamouring to be involved in these projects?

One remaining question to ponder – it is likely that if you belong to a public library then you already have access to online electronic books, whether they be reference works or technical books – so, how often have you, or do you make use of them and have you ever read one cover-to-cover onscreen?

Further Reading

Googlizers vs. Resistors. Brian Kenney. *Library Journal*. 12/15/2004.
http://www.libraryjournal.com/article/CA485756

Google Partners with Oxford, Harvard & Others to Digitize Libraries. Gary Price. *Search Engine Watch*.
http://www.searchenginewatch.com/searchday/article.php/3447411

Google's Library Project: Questions, Questions, Questions. Barbara Quint. *NewsBreaks*. December 27, 2004.
http://www.infotoday.com/newsbreaks/nb041227-2.shtml

1. Google Checks Out Library Books. Google Press Center.
http://www.google.com/intl/en/press/pressrel/print_library.html

2. Top Libraries to go online in deal with Google. Bobbie Johnson. *The Guardian*, Wednesday December 15, 2004.
http://www.guardian.co.uk/print/0,3858,5085842-103690,00.html

3. Oxford-Google digitisation agreement. Press Release. http://www.bodley.ox.ac.uk/news/news58.htm

4. Stanford University and Google Mass Digitization Project: FAQ.
http://library.stanford.edu/about_sulair/special_projects/google_sulair_project_faq.html

5. NYPL Partners with Google to Make Books Available Online. Press Release.
http://www.nypl.org/press/google.cfm

6. FAQ: The University's Pilot Project with Google.
http://hul.harvard.edu/publications/041213faq.html

7. Google's Library Project: Questions, Questions, Questions. Barbara Quint. *NewsBreaks.* December 27, 2004.
http://www.infotoday.com/newsbreaks/nb041227-2.shtml

8. International Libraries and the Internet Archive collaborate to build Open-Access Text Archives.
http://www.archive.org/iathreads/post-view.php?id=25361

Intranets and Content Management

Column Editor: Martin White
Intranet Focus Ltd;
e-mail: martin.white@intranetfocus.com

Another ten intranet success factors

A couple of years ago I developed a set of ten intranet success factors that ended up being used as a give-away on the UKOLUG stand at Online Information 2003. In rebuilding the Intranet Focus web site (which should look different by the time you read this) I thought it was time to review and revise the success factors, so here is the 2005 version. The main difference is really that the initial set were for start-up situations, and this set are for more mature (in relative terms) intranets. Comments would be much appreciated.

1. Document the intranet strategy
   Develop and document an intranet strategy, based on a consideration of the requirements and balance of information/content, technology and governance. This strategy has to be aligned with the overall strategy of the organisation and other relevant applications, such as records and document management, client/contact management and digital asset management. An important element of the strategy is defining acceptable measures of performance. A marketing strategy also needs to be included.

2. Understand and support content contribution
   Information quality in an intranet is essential, and yet most organisations implement a distributed approach to content contribution without really understanding the issues around content management. Invariably intranet content contribution is not included in job descriptions or valued by managers, and as a result gets low priority. The old adage 'garbage in – garbage out' certainly applies to intranet content!

3. Develop and enforce standards, and encourage the adoption of good practice
   There should be standards for intranets, especially where the organisation has a number of different intranets within some form of common architecture. These might typically include the basic layout of pages, default arrangements for lists, and conformance to accessibility standards. There have to be enforceable sanctions for non-conformance. A standard without such enforcement is not a standard, just a dream. In addition there will be good practice that may vary from department to department and from time to time. This good practice should be collated and reviewed.

4. Manage metadata
   Content authors may not have the skills and expertise needed to add metadata. Metadata is not just about 'keywords'. Structural metadata describes the information architecture of the document; content metadata provides a way of identifying documents that may contain relevant subject information; descriptive metadata enables the type of document to be identified. And finally administrative metadata deals information such as the person and department owing the document, the date when the document would be checked for relevancy, and the language of the document. The effort involved in developing metadata schemes is always underestimated.

5. Provide effective search functionality
   Users have to trust the search functionality so that they feel totally confident that either they have found all relevant information, or that information is not on the intranet. Not being sure one way or the other is not an acceptable situation. Implementing a good search solution will depend on a robust metadata environment, understanding how users will search for specific types of content, and considering the extent to which the intranet search engine should also search other repositories, such as records and document management applications.

6. Provide access to the business environment
   Make sure that the intranet provides access to information from external sources, such as business and market information. It is easy to be so focused on the provision of internal information that access to external information is overlooked. In today's competitive business environment monitoring the external business environment is of the highest importance.

7. Undertake regular usability testing
   Usability testing should be carried out at all stages of the life of an intranet, and there should be a range of feedback channels to ensure that the content and the information architecture continue to meet the expectations of contributors, users and stakeholders. Carry out regular surveys to obtain success stories that can be used to market the benefits of the intranet, and also identify examples

of where the intranet has failed to meet expectations. Do not make the assumption that because employees know enough about the structure and operations of the company that an information architecture based on organisational structure is all that is required.

8. Support collaborative working
    Although in theory intranets are supposed to support collaborative working little attention is usually paid as to what the real needs of the organisation are for collaborative working. It could be the support of project teams, of communities or practice, of extranet relationships with external organisations or for the assembly of complex, multi-author documents. Each of these may require other web technologies to be integrated with the intranet, such as wikis, blogs, instant messaging and discussion rooms.

9. Use workflow only where there is business benefit
    All too often companies implement workflow to try to control inappropriate content being publishing onto an intranet. Such breaches of personal responsibility should be dealt with through other channels. In implementing workflow little attention is usually paid to the impact on the email in-boxes, time available, and subject expertise of the reviewer/approver.

10. Quantify content migration issues
    Whether it is the large-scale migration of content into a CMS or the revision of sections of an intranet, the work involved is invariably underestimated. There is no substitute for a comprehensive content audit to understand what does need to be moved, and how, and what is now Redundant, Obsolete or Trivial (ROT).

Reference Management

Column Editor: Tracy Kent;
e-mail: t.k.kent@bham.ac.uk

Capturing notes and quotes

A common complaint about reference software tools is that they do not deal adequately with notes and quotes beyond adding keywords or abstracts. Some of the leaders in the market which do help in this area are provided below – purely in alphabetical order – which can be utilised with reference software

* *Devonthink* is a notepad, outliner, scrapbook manager, information manager, freeform database, archive, bookmark manager and image database with options for classifying and searching different reference types. http://www.devon-technologies.com/products/devonthink/overview.php

* *FileMaker Pro* which allows you to cross reference notes relationally with your bibliography and includes the ability to load the database onto the web. http://www.filemaker.com/products/fm_home.html

* *Hog Bay notebook*, for Apple's new Mac OS X operating system, lets you record and organize your ideas and relate concepts together. It is also scalable with numerous output options. http://www.hogbaysoftware.com/products/hog_bay_notebook.php

* *Knoesis* offers three products to help manage the note taking and reference process including the lite and professional index. The professional card index has the ability to share bibliographic information with programs like EndNote http://www.knoesis.com

* *Notelens*, for Windows, is a note taking tool which also integrates with Microsoft Outlook. It is based on free forms and allow the organising of your notes http://www.windsorinterfaces.com/notelens.shtml

* *Scholars Aid* keeps notes and references together and allows interaction with Z39.50 sites. http://www.scholarsaid.com/intro.html

Bookmark Management systems

There are many companies around which will manage your bookmarks but here is one which is free to join.

* *Connotea* is a bookmark management system which keeps links to articles read and the websites used so that you can find them again and are available from any terminal. It also collects the bibliographic information for the article or book that is being linked to from certain web sites. These include PubMed and Amazon. You can also add comments to your references to make them more meaningful. These references can then be exported into Endnote or other packages for future use. http://www.connotea.org

Research Support Tool

On a slightly different note is the *Open Journal Systems* which is a journal management and publishing system which seeks to improve access to research through a Public Knowledge Project. It assists with every stage of the referred publishing process from submission to online publication. In particular it has a

* Research Support Tool which link the study researchers are consulting to related research, Web sites, and databases. The website provides a demonstration which is quite useful. http://pkp.ubc.ca/demos/rsttour/index.html

Evaluation Templates

* Bibliography formatting software: An Evaluation template is now in its11th edition available at: http://www.burioni.it/forum/ors-bfs/text/index.html

Updates on Reference Software packages

* *Resource Mate* has made improvements to their demonstration version which include a makeover of their training booklet making it easier to read. The tutorial is available in html or pdf format and can be accessed as Word or RTF.
http://www.resourcemate.com

Bibliography management web services

* A current project of interest to readers of this column is that of *BibShare*, a set of bibliography management web services designed to be used by different text processing systems avoiding the need for different bibliography collections and formats. This is an example of the new generation of applications that have been built using a Service Oriented approach and is free

Current solutions to dealing with different bibliographies are limited in the sense that they are word-processor oriented, whereas users (often in Science and Technical areas) often write documents with different tools. This obligates authors to use different bibliography managers, and even different collections (with subsequent problems such as consistency enforcement, updating, etc.) depending on the word processor used.

Bibshare allows writers to use the same bibliography collection(s) regardless of the word processing system they use. Both personal and external collections can be used to retrieve the bibliographic information to be inserted into documents [original abstract amended]. For more information go to http://www.bibshare.org

Public Sector News

Column Editor: Jane Inman,
Warwickshire County Council and Chair of the
Affiliation of Local Government Information
Specialists (ALGIS in LARIA).
E-mail: janeinman@warwickshire.gov.uk

It is still e-government and Freedom of Information which dominate the public sector at present. Both aim to provide better information and services to the citizen and both are accepted as a good thing. It is important however, that we keep an eye on the resources these two areas can absorb and check that they are being used to best effect. As we do this we will have an eye too on the Gershon review of public sector efficiency which requires us to make 2.5% savings a year up to 2008.

E-Government

This is the year by which we in the public sector are required to be 100% e-enabled. Will we make it? Some reports being published suggest we will and equally as many who predict failure.

The most eagerly anticipated annual review of local government web sites conducted by SOCITM (The Society of Information Technology Management) was published in early March and shows 30% of local authority web sites have moved up a category. (Better Connected – 2005 http://www.socitm.gov.uk.) This report suggests local authorities may struggle with the Priority Service Outcomes set by ODPM last year.

Included in the Priority Service Outcomes are encouraging take up of services. A survey commissioned by the e-citizen National Project suggests that 17.5 million people could be in line to use local government e-services but that many are not aware of them. A campaign to improve take-up is essential. http://www.e-citizen.gov.uk will give you more information and the full report is due for publication in April 2005.

To steer us through this year there is a new face at the Cabinet Office. With the move of Ruth Kelly to Education after the Cabinet reshuffle David Milliband was given the responsibility for e-government along with public sector and civil service reform.

IPSV

I know it sounds rather like an unpleasant disease but the IPSV is actually the Integrated Public Sector Vocabulary and has just been made available in draft form. You can read more about it and view the draft at http://www.esd-toolkit.org

This vocabulary is a merging of the Government Category List with the Local Government Service List and the seamlessUK thesaurus and is designed to aid the indexing of public sector web sites. It will be used in the metadata of web sites and should assist joint working and data sharing as well as providing consistent indexing of local government web sites making them easier to search for the public.

Transport information

In December 2004 the Department for Transport launched the Transport Direct website which is intended to be the beginning of much more transport information delivered in a variety of ways. It is linked to 'Directgov' (http://www.direct.gov.uk) to provide, for instance, transport information alongside hospital or school details. It offers an integrated journey planner to find the fastest ways of traveling by public transport and will eventually be offered on the web, digital TV and mobile devices.

Health information

Health information is now being offered though a digital TV service launched by the Department of Health. http://www.dh.gov.uk has more information about this service known as NHS Direct Interactive. The existing NHS Direct web site receives about 6.5 million visits a year and they take 6.5 million phone calls.

Freedom of Information

We are now nearly 2 months into the new open culture of the Freedom of Information 2000. If my own experience is typical we have much to learn and to unravel. Some high profile examples of the use made of the Act have hit the headlines but I suspect the real cost is being felt in the less news worthy but still time consuming questions. In local government we have as anticipated received requests which although not difficult have been time consuming and we have found the use of the exemptions complex at times. Examples in my own area include requests for information about the number of Belisha beacons, details of buses, maintenance contracts and lots about speed cameras.

It would be nice to think it will settle down but experience in other countries suggests it will get busier before it settles but that there will always be high demand.

In central government the beginning of the Freedom of Information Act sees the end of the 30 year rule for the National Archives and they began the year by releasing more than 50,000 files which were less than 30 years old. Lord Falconer, the Lord Chancellor announced that central government had received 4,000 requests in the first month of the Act.

IDOX conducted a survey of local authorities calling it 'One month on-UK Local Authority responses to an FoI request' and 86% of the 438 authorities approached responded within the 20 days required by the Act. Rather impressive though was that responses were dispatched by no less than 220 unique job titles from Chief Executive to e-government programme manager!

Land Registry

Want to be nosey and find out what the house next door sold for? Well, the Land Registry online service can give you that information (if the property has changed hands since 2000) and all for just £2. http://www.landreg.gov.uk has been piloted for a while but has now been formally launched and is an ideal tool for identifying ownership of properties and prices paid.

With all the concerns about Data Protection it seems strange that this information is being made available so easily but the Land Register has always been a public document. All that has changed is that it is available at the desk top so you can find information it is interesting to know where as before you had to have a very good reason for travelling to check a Land Registry entry.

Museums, Libraries and Archives Council (MLA)

The MLA People's Network is looking for librarians to join its Online Enquiry Service. This service will aims to encourage all 149 library authorities in England to get involved and provide the public with another way of accessing quality information. For more information go to http://www.peoplesnetwork.gov.uk

Government websites

The Economic and Social Research Council have produced a report which gives some interesting figures for those of us involved in e-government and specifically the web based provision of information.

The report shows 40% of internet users visiting news and current affairs web sites but only 28% had used local council information and 21% central government. While only 2% had visited the web site of their own MP, research shows that 44% of the public would like their MP to publish an email address and 42% would like them to have a web site. You can read more about the findings on the ePolitix web site http://www.ePolitix.com

You may find useful the Public Bodies Database which lists the public bodies sponsored by central government and gives contact details, information about their remit and board members. It may be found at http://www.knowledgenetwork.gov.uk/ndpb/ndpb.nsf/0/04801FA2F8BF454480256E540049FF84?OpenDocument

Public Sector News is provided by the ALGIS (The Affiliation of Local Government Information Specialists) which represents information professionals providing information services to local authority staff and elected members. For more information go to http://www.algis.org.uk

Current Awareness

Column editor: Jane Grogan
e-mail: Jane.Grogan@gchq.gsi.gov.uk

This column contains summaries of articles (print and electronic) about online services, CD-ROMs, networked information, electronic publishing, multimedia etc. including, with permission, abstracts identified with an * next to the author initials, drawn from Current Cites, the monthly publication distributed electronically by the Library, University of California at Berkeley:
http://sunsite.Berkeley.edu/CurrentCites/.

The Current Cites database of citations now numbers over 1,000 and is searchable at http://sunsite.berkeley.edu/CurrentCites/bibondemand.cgi. This service is called "Bibliography on Demand" because it can be easily used to create printable or linkable bibliographies on information technology topics.

EDUCATION

Meadows, Jack. "The immediacy effect – then and now" *Journal of Documentation* 60 (6) (2004): 601-608. Provides an overview of the bibliometric studies that have examined the rate of obsolescence among disciplines, or between the different fields within a discipline. Discusses some of the assumptions made, and sets out some ideas for new questions to be answered for the librarian in charge of electronic library collections. - [CJU]

Wilder, Stanley. "Information Literacy Makes All the Wrong Assumptions" *The Chronicle of Higher Education* 51(18)(7 January 2005): B13. (http://chronicle.com/prm/weekly/v51/i18/18b01301.htm). - Wilder makes an interesting case for why information literacy programs, a staple at many academic libraries, are a bad idea. Wilder cites such problems as students who feel like they don't need any help finding information, and our inability to reach even a fraction of the potential audience with formal instruction. After arguing his case, he then suggests an alternative model for librarians as teachers, stating that "Librarians should use their expertise to deepen students' understanding of the disciplines they study." By pairing discipline-based and situational teaching moments with smarter, more effective and more easily used information systems, "we can create educational programs that reach everyone on our campuses, every time they turn to us." [full disclosure: I'm quoted in the piece] – [*RT]

ELECTRONIC PUBLISHING

Appleton, Leo. "The use of electronic books in midwifery education: the student perspective" *Health Information and Libraries Journal* 21 (4): 245-252. Describes the evaluation of a pilot implementation of an e-book collection for student midwives in a college of higher education. Focus group (n=10) findings showed that students viewed e-books as alternatives to printed books, considered payment of use of e-books entirely unfair, and did not like slow downloading (particularly when studying remotely). They liked the idea of having chapters available via a VLE, and would like to search full text for information they required, to obtain chapters, in much the same way as they search for journal articles. - [CJU]

Gnatek, Tim. "Libraries Reach Out, Online" *The New York Times* (9 December 2004) (http://www.nytimes.com/2004/12/09/technology/circuits/09libr.html?ex=1260248400&en=bc31f3ce53fcf024&ei=5090&partner=rssuserland). - Yeah, yeah...we know. E-books are dead. Aren't they? Apparently not, as public libraries add extensive collections of electronic books, "laying claim to a massive online public as their newest service audience." This article discusses the New York Public Library's new collection of 3,000 electronic titles and how check-in and check-out by websurfing cardholders is handled. And the NYPL is not the only system offering e-books to its customers. The article mentions similar collections at the Cleveland Public Library and the King County Public Library. Libraries are also beginning to offer audio books in downloadable MP3 format, and some systems are even putting movie trailers online. The article also mentioned online book clubs, virtual reference, IM-based tutoring, library-sponsored LAN parties for online video gaming and free wireless Internet hotspots. – [*SK]

McGrath, Eileen L., Metz, Winifred Fordham, & Rutledge, John B. H-Net Book Reviews: Enhancing Scholarly Communication with Technology. *College & Research Libraries* Vol 66 (1) January 2005: 8 – 19. This article describes the evolution of free, scholarly electronic book reviews which are available from the electronic forum H-Net: Humanities and Social Sciences Online. H-Net is an umbrella organisation for approximately one hundred and thirty-eight listservs in the Humanities and Social Sciences subject areas which has been growing since 1992. A comparison is made between electronic-only reviews available via one listserv, H-South, and the same books which are reviewed in an established print journal, The Journal of Southern History (JoSH). The study successfully answers the question are the electronic-only reviews published via H-South significantly different from the print reviews found in JoSH? The conclusion is yes. Criteria used to answer this question are: the choice of books for review, the academic status of reviewers, how positive or negative evaluations are and an assessment of how evaluative or descriptive are the reviews. The authors draw attention to the increased availability of H-Net Reviews because of the decision by the Library of Congress to link records in its online catalogue to H-Net Reviews. H-Net Reviews are easy to retrieve through Google by typing "H-Net" and the title of the book. [AS]

GENERAL

Ball, Mary Alice. "Libraries and University Presses Can Collaborate to Improve Scholarly Communication or "Why Can't We All Just Get Along?"" *First Monday* 9(12) (6 December 2004) (http://www.firstmonday.org/issues/issue9_12/ball/). - Ball has written a succinct but thorough overview of the historical underpinnings of "scholarly communication," where we are now, and where we are going. She approaches this high profile topic with an eye to the competing cultures of book publishing and the library profession. Both have much to learn from each other, she argues, and both have much to gain. Publishers maintain dramatically high standards for final products, and they are powerhouses for design, marketing and selection strategies. Librarians currently hold a clear edge in grasping the importance of information standards, "fair use" as a social good, and the power of the library's imprimatur. Ball is looking for the common ground between the two, but she also points out that the fault line between publishers and librarians is copyright compliance. She argues that the two camps should develop common goals in educating and influencing university administrations and faculties -- and that at the present moment, both groups have an opportunity to be heard and understood. This is an excellent article not only for its historical analysis, but also because it does a superb job of defining the motivations that drive publishers and librarians. This leaves the reader with an overview that is a fertile ground for brainstorming. – [*TH]

Dudman, Jane A global adaptation *Information World Review*, Issue 210, February 2005. Pages 16 - 17 Interview with Richard Charkin, CEO of Macmillan Publishing, talking about the company's strategy. Macmillan's business overseas is now much larger than in the UK. Charkin talks about the challenges of the Internet, the company's use of technology and the challenge of copyright. – [LR]

Entlich, Richard. "One Last Spin: Floppy Disks Head Toward Retirement" RLG *DigiNews* 8(6) (15 December 2004) (http://www.rlg.org/en/page.php?Page_ID=20492&Printable=1&Article_ID=1692). - This informative piece summarizes the history of the floppy disk in its various incarnations, as well as highlighting why the format has endured as long as it has. The reasons why data that exists only on floppy disks is endangered are outlined, as are strategies for rescuing the data. It's clear that floppy disks are on their way out, we just don't know exactly how fast. But as this piece makes clear, we would be wise to start getting the data off those disks as soon as we can, since it will only get more costly and/or more difficult the longer we wait. – [*RT]

Flood, Gary Tame the mountain *Information World Review,* Issue, Issue 210, February 2005. Pages18-19 Article giving tips on email management and some technical solutions. – [LR]

Greenfield, Adam. "All watched over by machines of loving grace: Some ethical guidelines for user experience in ubiquitous-computing settings " *Boxes and Arrows* (23 November 2004) (http://tinyurl.com/6s87p). - What if we had computers embedded in every little gizmo in our waking lives and their design was no better than the average voice mail system and their privacy guidelines looked like they were put together by a telemarketer? It'd be hell. These two concerns, design and privacy, are the focus of this article on Ubiquitous Computing or "ubicomp" by Adam Greenfield. In it, Greenfield paints a nightmare world of ubiquitous interruptions and widgets so unusable that they spin out-of-control at the slightest slip of the finger. Greenfield goes on to propose a number of "baseline standards" for how systems ought to work from the standpoint of systems and interface designers. Several of the comments following the article are also worth reading, particularly the heartfelt one by Chris Fahey on Opt-in v. Opt-out. – [*LRK]

Hiller , Steve. "Measure by Measure: Assessing the Viability of the Physical Library" *Bottom Line* 17(4) (2004): 126-131. - It's never been part of my job to close a branch library, or to open one for that matter. Part of (surviving) the process however must surely consist in establishing criteria that are acceptable to most, if not all, whether the eventual decision is thumbs up or thumbs down. In this article, we get a little background on the evolution of how they make these decisions at the University of Washington in Seattle. The criteria they use has undergone an evolution even in the past few years as the implications of remote access and changing user study habits become clearer. The author describes four broad categories such as "use" and "facility quality" that they use to evaluate branch library viability. These categories in their turn are broken down into five different measures. Given this metric, the surviving branch libraries, in the words of the author, "will be those that are still dependent on print collections and that provide space that supports student work in a collaborative teaching and learning environment." – [*LRK]

Kilkki, Kalevi. "Sensible Principles for New Networks and Services" *First Monday* 10(1)(3 January 2005) (http://www.firstmonday.org/issues/issue10_1/kilkki/). - The author reviews the development and history of network design over the years since it really hit big, and offers three relatively simple -- but very powerful -- suggestions on how to approach design for the long-term. When you get to the conclusion and read them, you might well say, "Knock me over with a feather" --because they sound like common sense solutions than anyone would think of. In fact, we have learned that network environments, which use many complex technologies that increasingly depend upon one another in cascading relationships, have a tendency to lead designers away from common sense solutions by nature of their very complexity. Kilkki analyzes the reasons for past design failures as a preface to his three-point plan. First he argues that the analysis of customer needs must focus on practical uses that stand a good chance of becoming everyday routines. Second, a new technology or product should be based on well-defined, carefully selected core principles. Third, the real experiences of users must be taken into account continuously, not just at the outset and the nadir of a product lifespan. Obvious stuff, right? Well, the Internet is littered with well-intentioned "bloatware" that fails one or all three of these design principles, so this article is good reading indeed. – [*TH]

McCartin, Shona and Reid, David. D-I-Y Interloans: from dream to reality *Electronic Library* 22(6) 2004 509-517 This paper describes a pilot project that allowed end-users at Lincoln University Library, Canterbury, New Zealand, to initiate their own inter-library loans using the local implementation of VDX (Voyager). Initially the implementation of VDX within New Zealand had excluded the end-user request interface. Trends and growth patterns in interlending, how the end-users coped, customer reactions, and benefits for the library staff are described, and an indication is given of future trends. While it may be of interest to interlending specialists, this article does not convey a particularly clear picture of the issues, and is actually quite difficult to follow for someone not versed in the details of VDX. – [CME]

Seffah, Ahmed, and Eduard Metzker. "The Obstacles and Myths of Usability and Software Engineering. " *Communications of the ACM* 47(12) (December 2004): 71-. - Usability and User-Centered Design (UCD) permeate so much of what we do on the web. It sometimes comes as a shock to find out that the world of programming may be marching to an entirely different beat. In fact, as this article seems to suggest, there is a considerable disconnect between the worlds of interface designers and programmers. The article speaks to both groups, assigning blame in a most even-handed and perceptive way. It also suggests sensible ways of bridging the gap. These include rationalizing UCD methods as well as making these methods a standard part of programming/computer science curricula. – [*LRK]

Town, Stephen. "View E-measures: a Comprehensive Waste of Time?" *Vine* 34(4): 190-195. - This is the kind of article I love reading. In Vine's special issue devoted to "Library evaluation in practice", the author argues that maintaining stats on the use of e-resources ("e-measures" as he calls them) is a "waste of time". Once past this self-consciously provocative statement, the author makes clear that the numbers only have meaning when their context is fully understood. "Counting," he declares, "is still no substitute for listening." In this respect, he has praise for methods like LibQUAL+ which make satisfaction and experience measures a vital part of overall evaluation. "What we decide to measure in the library e-environment," he warns, "will result in adaptation of our behaviour to maximise the chosen measures. In so doing it will help define what future libraries are and what they do, and demonstrate their priorities." – [*LRK]

VRD 2004 Online Proceedings Syracuse, NY: Virtual Reference Desk Project, November 2004. (http://www.vrd2004.org/proceedings/). - These are about three dozen of the presentations given at the 2004 Virtual Reference Desk Conference in Cincinnati, OH, November 8-9, 2004. Divided into sections on Management, General, Evaluation and Standards, Technology, Resources, Research and Policy, and Vendor Demonstrations, they are mostly MS PowerPoint slides converted to an HTML presentation. Unfortunately, choices made during this conversion provide unreasonably small screen sizes, which renders screenshots of

browser windows (a frequent component of virtual reference presentations) virtually unreadable. Such a decision is inexplicable in the age of broadband, but complaining is unlikely to change anything. Meanwhile, there is a good deal of information that can be extracted from the other slides and there is probably no better way to feel the pulse of virtual reference in libraries today. – [*RT]

INFORMATION ACCESS/RETRIEVAL

Chillingworth, Mark Shaken not stirred *Information World Review*, Issue 210 February 2005. Pages 13-14 Interview with Mike Lynch, CEO of Autonomy. He discusses his strategy to foil Google and their plans to move into enterprise search. – [LR]

Chudnov, Daniel, and Jeremy Frumkin. Service Autodiscovery for Rapid Information Movement (10 December 2004) (http://curtis.med.yale.edu/dchud/writings/sa4rim.html). - This paper explores issues relating to capturing resource citations and links, routing them to various locations, and using them with link resolvers and other services such as "gather, create, share" tools. And that's just for starters. They quickly move into discovery autodiscovery -- first link autodiscovery, then metadata autodiscovery, and finally service autodiscovery. There is a lot to absorb in this paper, but it's well worth spending the time to absorb it. They are doing no less than rethinking how we both offer our services as well as consume the services of others, based on a brave new world chock-full of new and powerful web-based applications and services. I can think of no better example of the kind of imaginative thinking we nee d to make effective use of our opportunities and challenges. – [*RT]

Douyère, Magaly, Soualmia, Lina F., Névéol, Aurélie, Rogozan, Alexandrina, Dahamna, Badisse, Leroy, Jean-Philippe, Thirion, Benoît, and Darmoni, Stefan J. "Enhancing the MeSH thesaurus to retrieve French online health resources in a quality-controlled gateway" *Health Information and Libraries Journal* 21 (4): 253-261. Describes how the MeSH thesaurus was expanded to enable it to be adapted for use in the CISMeF, the Catalogue and Index of Health Resources in French. Work included addition of two new concepts to deal with Internet resources: resource types (similar to the publication types) and metaterms (generally a medical specialty or a biological science). Illustrates a typical query expansion process, and discusses some of the possible navigation features required. - [CJU]

Fallows, Deborah. "Search Engine Users" *Pew Internet & American Life Project* (23 January 2005) (http://www.pewinternet.org/pdfs/PIP_Searchengine_users.pdf). - This latest study in the Pew Internet & American Life Project asked people about their search engine attitudes and habits. There is no shortage of contradictions in the findings. Among other things, the study found that just one in six search engine users could tell the difference between unbiased search results and "sponsored" search results. In general, users also had little understanding of how search engines work. Nonetheless, some 92 percent of search engine users expressed confidence in their searching abilities. And -- almost half of the respondents said they would actually stop using search engines if these sites were not clear about how the paid results were displayed. Go figure. Interestingly, some two-thirds of those surveyed said "they could walk away from search engines without upsetting their lives very much." Whereupon they would "return to the traditional ways of finding information." Hmmm... Use the library, perhaps? Call a friend or relative? Approach a colleague in the office break room? Or maybe they would just blow it off altogether. The study found that 33% of searchers "would not bother looking up most or even all of the information they search for if they lacked access to internet search engines." – [*SK]

Ferguson, Charles H.. "What's Next for Google" *Technology Review* 108(1)(January 2005): 38-46. - We've heard so much about it lately, the question is more like what isn't next for Google. I'm sure they're hoping that butting heads with Microsoft isn't. However, Charles Ferguson argues that both giants' current development of uber-search systems for search & retrieval from all sources of digital information (the public web, the "dark" web, your intranet, your hard drive, and on) makes a collision highly likely. Part explanation of the technology, part analysis and advice about business strategy, the article resonates with the experience of someone who's walked with the giants and avoided getting stomped: Ferguson co-founded Vermeer Technologies, which released the FrontPage website development application and immediately faced competing technology from Netscape and Microsoft. (He sold to Microsoft for a nice big number.) Those of us who care about control of vital markets will find this quite interesting; those who don't can still use this preview of the tools in our future. Information providers, prepare to adjust yet again. – [*JR]

Hahn, Karla L., & Schmidt Kari Web Communications and Collections Outreach to Faculty. *College & Research Libraries* Vol 66 (1) January 2005: 28 – 37. This article describes a qualitative survey of the websites of Scholarly Publishing and Academic Resources Coalition (SPARC) libraries to discover how often they are using the Web to communicate about collection changes and scholarly communication issues with their faculty and researchers. The methodology involves an unobtrusive survey analysing the content of one-hundred and forty-nine research and college libraries' websites in the United States and Canada. As the study population are all full members of SPARC it is assumed that they have an active interest in educating faculty and researchers about the crisis in scholarly communication and the impact this has upon the local collection. Findings are disappointing and indicate that web pages are infrequently used to present information on scholarly information or the specifics of collection management projects. Instead static information such as general collection overviews, request forms and lists of collection contacts tend to be provided. Eight libraries are identified as models of good practice with webpages that discuss collection changes in detail, as well as scholarly communication issues. - [AMHS]

Sihvonen, Anne and Vakkari, Pertti. "Subject knowledge improves interactive query expansion assisted by a thesaurus" *Journal of Documentation* 60 (6) (2004): 673-690. Compared the searches conducted by experts (15 undergraduate education students) and novices (15 other undergraduates). The search condition was realistic to the extent that the students were using these in their coursework assignments. Two tasks (one easy, one difficult) were selected, and students were supported in their searches by suggestions from the ERIC thesaurus. Students were asked to think aloud in the sessions and their comments recorded. In addition, search logs were analysed. Searching patterns varied, with experts using more of their own terms, as well as thesaurus terms in the expansion terms. One significant difference was the greater use of 'broader terms' by the experts. The experts performed better than the novices on the difficult task, and experts could use the thesaurus to improve search effectiveness, unlike the novices. Concludes that familiarity with the topic area is necessary to benefit from using a thesaurus. - [CJU]

METADATA

Bittern, Diana. Where's my link? Why linking is still not perfect. *Managing Information*, 11 (10) (December 2004): 46-48. Many institutions invest significant resources into purchasing electronic full text resources and then spend more to purchase and implement link resolution systems that promise to connect end users directly from a bibliographic citation to the full text of the article. This paper looks at some of the realities of linking and the steps being taken by vendors to overcome anomalies in matching citations from bibliographic databases to the full text. Linking is essentially about matching the metadata to the article, but there is no standard for the presentation of metadata. The most common metadata matching information is IVIP or ISSN, Volume, Issue, Page. This is considered preferable to other metadata elements, such as title and author, because it reduces the possibility of variations. Even so, there are still myriad variations, such as combined issues and supplements, to contend with. Even the ISSN, considered the unique identifier for a serial, can be problematic. Some publishers have different ISSNs for the print and electronic versions of the same journal, some do not. Software providers have begun to develop the means for linking across versions, but it is far from ideal. The CrossRef project was set up by publishers to promote cited reference linking by compiling a database of persistent identifiers with relevant metadata. In theory it may be "the silver bullet," but in practice it can suffer from incomplete data, data inconsistencies and variations in technical capabilities among the member publishers. - [LR]

Descriptive Metadata Guidelines for RLG Cultural Materials Menlo Park, CA: Research Libraries Group, January 2005.(http://www.rlg.org/en/page.php?Page_ID=214). - Although this document comes out of RLG's need to describe metadata requirements for participants in such RLG projects as Trove.net and RLG Cultural Materials, they can inform and be used by a wide range of projects that depend on good metadata. The sixty-seven page documents begins with defining terminology, then moves on to guidelines for data fields and structure, data content and values, data formats, and core descriptive fields. It would be unwise to assume that the section on data conversion for RLG Cultural Materials can be skipped if your institution is not a contributor, as this section is a fascinating look at the kinds of metadata transformations that are required to create a sensible union catalog. This document is chock-full of excellent advice, useful examples, and hard-won metadata wisdom. It should be required reading for anyone working with metadata. – [*RT]

Godby, Carol Jean, Jeffrey A. Young, and Eric Childress. "A Repository of Metadata Crosswalks" *D-Lib Magazine* 10(12) (December 2004) (http://www.dlib.org/dlib/december04/godby/12godby.html). - Librarians must increasingly deal with metadata in a wide variety of formats. "Dealing" with such formats will, in many cases, mean transforming the data from one format to another. The main mechanism for doing this is a metadata "crosswalk" or a specification as to how one format can be translated into another. Since such a

procedure will hopefully be done with software rather than humans, it is necessary to specify a machine-parseable mechanism to handle crosswalking. This article specifies one such piece, and a pivotal one, if we are to enable computers to seamlessly move metadata around. Not surprisingly this work is being done by OCLC Research, an office that clearly both "gets it" and has the technical wherewithal to do something about it. We would be wise to pay attention. – [*RT]

Haas, Stephanie C.. "X Marks the Spot: The Role of Geographic Location in Metadata Schemas and Digital Collections" *RLG DigiNews* 8(6) (15 December 2004) (http://www.rlg.org/en/page.php?Page_ID=20492#article1). - This is a good introductory/overview article on geographic location information for digital objects. Starting with how MARC and LCSH deal with this information, Haas moves on to more recent and sophisticated examples of geospatial data in modern computer systems. Examples of such systems include the National Oceanic and Atmospheric Administration library catalog, the Alexandria Digital Library, and the Electronic Cultural Atlas Initiative. The sources cited can serve as a start for further investigation. – [*RT]

Van de Sompel, Herbert, Michael L. Nelson, and Carl Lagoze, et. al. "Resource Harvesting within the OAI-PMH Framework" *D-Lib Magazine* 10(12) (December 2004) (http://www.dlib.org/dlib/december04/vandesompel/12vandesompel.html). - The Open Archives Initiative - Protocol for Metadata Harvesting (OAI-PMH) is a well-established protocol for retrieving piles of metadata from compliant content repositories. One of the most well-known harvesters is OAIster, which has gathered records for nearly 5 million digital objects from OAI-compliant repositories. This article, by those directly involved with the development of the protocol, looks at how service providers (those who gather metadata from data providers) can use the protocol to fetch the actual objects described by those metadata records. After reviewing possibilities relating to protocol extensions, they discard those options in favor of specifying a metadata format that can better accommodate resource harvesting than the protocol-required simple Dublin Core format. Since the protocol already specifies that any metadata format can be surfaced in addition to simple Dublin Core, no protocol extension is necessary to enable content harvesting. The authors advocate the use of the MPEG-DIDL format, although they acknowledge that the METS format could be used for this purpose as well. – [*RT]

PRESERVATION

Foster, Nancy Fried, and Susan Gibbons. Understanding Faculty to Improve Content Recruitment for Institutional Repositories *D-Lib Magazine* 11(1)(January 2005) (http://www.dlib.org/dlib/january05/foster/01foster.html). - I've become a fan of the University of Rochester. Well, not the university itself, really, but the library, where they really seem to get it. This article is no exception, as through studying faculty and what is important to them they determine how they can best influence faculty adoption rates of institutional repositories (IRs). What they find is that "It is essential that anything in an IR be absolutely safe and secure. Beyond that, the single most important criterion of an IR's value to our faculty members is that other people find, use, and cite the work that they put into it. Even the most enthusiastic supporters of IRs will soon lose interest if this criterion is not met." The authors then go on to describe how they added another component to their DSpace code to provide personalized views of IR content. However, the key take away here is how they learned to better pitch the IR to faculty -- a lesson that I'm guessing a number of IR projects still need to learn. – [*RT]

Holley, Rose. Developing a digitisation framework for your organisation *Electronic Library* 22(6) 2004 518-522 This paper discusses the elements of a framework for digitisation activities at the University of Auckland Library, how it was developed, and what has been learned from the work. The style is very compressed – the text seems to consist largely of bullet points – with extensive references to other documents. Despite this, it is an exceptionally clearly-written and useful paper, and would serve as an excellent introduction to digitisation issues for anyone embarking on digital projects. The author starts out by giving a brief account of the global and local New Zealand pictures, then progresses to a detailed account of her work and an outline of proposed future developments. She concludes with a short section on 'developing a digitisation framework for your organisation'. – [CME]

Kenney, Brian. "Googlizers vs. Resistors" *Library Journal* (15 December 2004) (http://www.libraryjournal.com/article/CA485756). - This article is a summary transcript of a debate-like discussion held at the Pennsylvania Library Association Annual Conference in October. Although it predated the announcement by Google that they were preparing to digitize the full collections of selected major research libraries, there was nonetheless enough fodder for a spirited discussion of issues. I doubt anyone will be "converted" from a given perspective to another from reading this piece, but that is not its purpose. If the piece causes us to think about our services in relation to the services of web sites like Google, and consider carefully our appropriate role, then it will have been well worth the reading. – [*RT]

Quint, Barbara. "Google's Library Project: Questions, Questions, Questions" *Information Today NewsBreaks and the Weekly News Digest* (27 December 2004)(http://www.infotoday.com/newsbreaks/nb041227-2.shtml). - Here's an article that you may have missed in the post-Christmas afterglow. Barbara Quint asks a number of questions about Google's gargantuan cooperative digitization project with selected research libraries, and she gets answers from diverse individuals. Some of the most interesting responses are to the question: "What impact could this project have on current digitization projects?" The manager of a digital library project that provides access to over 10 million images says that his and all other digitization projects have suddenly become "small-scale." A research library consortium spokesperson "predicted that the new project could table or even kill current digitization projects at libraries, while the librarians waited to see if their planned projects were necessary or, assuming their content was unique, if Google might someday digitize that content for free." (An unexplored issue is the impact of the Google project on funding agencies' interest in future digitization projects by non-participating libraries.) Can you say paradigm shift? John Berry thinks it's one. If you're feeling a bit queasy from that shift, don't forget that the Internet Archive and ten libraries from around the globe announced shortly after the Google revelation that they would digitize over one million books. – [*CB]

SECURITY

Electronic Privacy Information Center. "2004 Privacy Year in Review" *EPIC Alert* (11 January 2005) (http://www.epic.org/alert/EPIC_Alert_yir2004.html). - The Electronic Privacy Information Center has compiled the "Top Ten Privacy Stories of 2004". Read 'em and weep: Foreign Opposition to USA PATRIOT Act; Google Datamines Private Email; Expansion of US-VISIT ("entry-exit border control system"); Death of Airline Passenger Profiling...Maybe; U.S. Medical Records Go Overseas (outsourcing); Data Disclosures--Mission Creep Continues (the IRS, the Census Bureau...); States Pull Out of Mini-Total Information Awareness Project; ID Theft a Growing Problem; Laws Stiffen Penalties; Prevent More Stringent ID Requirements for Voters; California Continues Privacy Reforms. These are important issues, and a paragraph-long discussion of each one provides a basic level of understanding. EPIC also identifies privacy issues to keep an eye on in the coming year, such as national ID cards; renewal of the USA PATRIOT Act; an attack on privacy rules by the telemarketing industry (involving the Do Not Call List); Google's potential for tracking what you read via its new library book digitization project; VoIP privacy issues; "Smart Barcodes, RFID, and Products that Spy"; Internet privacy (some things never change); outsourcing (your personal data now going offshore); centralized voter registration databases (technical expertise is lagging); and the WHOIS directory (which "still lacks basic privacy safeguards"). – [*SK]

O'Harrow, Robert Jr.. "ChoicePoint Finds Wealth in Information" *Washington Post* (via MSNBC) (20 January 2005)(http://www.msnbc.msn.com/id/6846357/). - Since 1997, ChoicePoint has morphed into "an all-purpose commercial source of personal information about Americans, with billions of details about their homes, cars, relatives, criminal records and other aspects of their lives." Not only is it awash in corporate and government clients, but it is rapidly becoming "a private intelligence service for national security and law enforcement tasks" -- acquiring companies that make sophisticated data mining tools along the way. Note: "ChoicePoint and other private companies increasingly occupy a special place in homeland security and crime-fighting efforts, in part because they can compile information and use it in ways government officials sometimes cannot because of privacy and information laws." The associate director of the Electronic Privacy Information Center said ChoicePoint is assisting in the creation of "a Scarlet Letter society," since it's virtually impossible now for anyone to shake off even minor past transgressions like bounced checks. ChoicePoint is not the only company in this business. "An entire industry has mushroomed during the past decade because of extraordinary increases in computing power, the expansion of telecommunications networks and the ability of companies like ChoicePoint to gather and make sense of public records, criminal histories and other electronic details that people now routinely leave behind." Read about their role in everything from national security to pre-employment screening at big box retailers. – [*SK]

WEB DESIGN

Hammond, Tony, Timo Hanay, and Ben Lund. "The Role of RSS in Science Publishing: Syndication and Annotation on the Web" *D-Lib Magazine*   10(12) (December 2004) (http://www.dlib.org/dlib/december04/hammond/12hammond.html). - "RSS is the very antithesis of the website," the authors -- from the Nature Publishing Group -- point out here, indicating that it functions more as a syndication/annotation tool. They offer an explanation of RSS technology and development, and indicate that it is catching on fast in the world of scientific publishing because it "presents a very simple XML structure for packaging news titles and links, and delivering them down to user desktops and handhelds." RSS, according to the authors, "allows us to bundle rich descriptive metadata" along with the standard newsfeed items. This alone would make RSS attractive to science publishers. It also serves as an excellent delivery vehicle for tables of content alerting services. From a purely scientific standpoint, RSS can be used "transmit complete scientific data sets." The authors discuss Urchin, an open source RSS Aggregator developed by the Nature Publishing Group. "NPG uses Urchin to provide keyword-filtered RSS feeds for its staff, and to populate a science, technology and publishing news portal." The article contains extensive notes and a bibliography. – [*SK]

Contributors to Current Cites * : Charles W. Bailey, Jr., Terry Huwe, Shirl Kennedy, Leo Robert Klein, Jim Ronningen, Roy Tennant

Contributors: David Bevington, Peter Chapman, Sarah Clarke, Catherine Ebenezer, Jonathan Gordon-Till, Dave Huddart, Linda Kerr, Marty Kesselman, Ida Kwan, Penny Leach, Shona McTavish, Shirley Parker-Munn, Liz Reiner, Lesley Robinson, John Ross, Ann Smith, Christine Urquhart, James Watson.

Meeting Review

Developing and Managing e-book Collections

John Rylands University Library, 2nd Feb 2005

It seems that everyone wants to know about e-books now. Perhaps we see them as the solution to multiple copy problems, or a way of resolving space issues or increasing access to our collections 24/7. I certainly have hopes that they will help with all these issues, but I have also been struggling with practical problems of finding out just what is available, how I can buy it and a myriad of technical problems. This workshop promised to help with some of these issues.

I was surprised that the majority of participants were not subject librarians, like myself, but came from acquisitions departments and technical support sections too. Perhaps this is an indication of how the move towards e-books is affecting processes right across the Library.

The workshop began with a definition. What exactly is an e-book? There are a variety of different descriptions available, but Chris Armstrong and Ray Lonsdale, the workshop facilitators, provided the following; " any piece of electronic text regardless of size or composition (a digital object), but excluding journal publications, made available electronically (or optically) for any device (hand held or desk bound) that includes a screen."

Armed with this useful explanation, we then had a practical opportunity to explore a range of e-book types: reference, fiction, textbooks and monographs. As we looked at the examples, we were asked to comment on the suitability of each resource for our own library and what we felt the advantages and disadvantages were. I particularly liked the Spartacus history textbook (http://www.spartacus.schoolnet.co.uk), in spite of the intrusive sponsored links and the range of texts available in NetLibrary (http://www.netlibrary.com). Conversely, I was concerned about the opportunities for plagiarism encouraged on the Literature Network site (http://www.online-literature.com).

The main focus of the day was a practical group exercise in which we discussed a set of issues relating to the development of e-book collections. This was an extremely useful way of sharing experiences and learning from both the facilitators and other participants. EBook Library and it's link with Dawsons was mentioned as a good means of identifying ebooks, but the group felt that a national bibliography of ebooks would be very useful. Problems relating to acquisition of ebooks included issues of licencing, cost, bundling, archiving, hardware, networking and many more. My group discussed the frustrations of sourcing a single textbook title rather than a whole collection. Facilitating

access, evaluating use, marketing and promotion were amongst the other topics covered and it was interesting to hear how other libraries are dealing with these issues.

We concluded the day by summarising what we felt the main advantages and disadvantages of e-books were at present. Advantages include space saving, no damage, 24/7 access, multiple access, off campus access, no shelving, integration with curriculum, value added features such as searching content and manipulation of data (http://www.knovel.com/knovel2/default.jsp) and tailoring texts.

Top of the disadvantages for my group was that publishers still don't understand library budgets, followed by the lack of availability of key texts, incompatibility of software and lack of library control, to mention just a few! The workshop was supported by a useful workbook which includes a lot of background reading and links to examples of e-books, which I have been able to share with my colleagues at work.

All in all, a very informative and enjoyable day.

Moira Bent
Faculty Liaison Librarian
Robinson Library
Newcastle University.

Book Reviews

The content management handbook
Martin White. London: Facet Publishing, 2005
ISBN 1-85604-533-1 176pp £39.95

If your organization is struggling with managing its intranet or internet website, then you have probably thought about getting a content management system. This book should be your starting point. It provides step-by-step guidance on how to develop a content management strategy, sets out the elements of a content management system and gives advice on the specification, selection, implementation and management of content management software.

The author's aims for the book are modest, and it is intended to be complementary to those well-regarded texts, Boiko's *Content Management Bible*, and Rockley's *Managing Enterprise Content*, by taking a project view of what it takes to get from wondering how implementing a CMS could benefit the organization through specifying, selecting and procuring a CMS product, to implementing the product on time and on budget. It draws on seminars held by UkeIG on content management.

It has three sections: information issues, technology options and governance issues. There are chapters on: creating an information-enabled organization; content management functionality; information architecture and

metadata; developing a content management strategy; technology options and costs; making the business case; the CMS business; managing the project; writing a statement of requirements; the selection process; implementation; content migration, and resources. I found the chapter on content management strategy particularly useful; the author notes that there has been a welcome and long-overdue shift from seeing a website as a repository of information that the organization thinks would be useful to a visitor to seeing a website through the eyes of a user. He covers the use of personas to develop an information architecture for a site.

It is gratifying that information professionals get a plug, but not surprising given its publisher. However, I did wonder that putting this plug at the beginning might put off other readers, e.g. IT or communications managers who might be leading CMS projects. In the introduction Martin White claims that there is "virtually no area … where information professionals cannot make a significant and on-going contribution to the effective management of content, and thus information and knowledge, in any organization that is struggling to cope with business in the 21st century." And it is true that information professionals have a role in metadata, search, information architecture, and information and content audits. It would have been a bonus to have case studies included, a feature of Orna's Practical Information Policies that I found particularly interesting. In this there is only one, on personas used by a university project. Martin White's approach represents the ideal; sometimes it is useful to learn lessons from projects where the ideal staffing or time was just not available.

There are a few signs of hasty editing and some gremlins, but these are minor and do not detract from the usefulness of this book. The annotated list of resources is especially useful. When there are so many websites and books on content management , it is helpful to focus on those recommended by an expert. The index is reliable. Each chapter has a summary of key points, handy for showing to project sponsors, for example. Throughout the text there are diagrams, summary tables, and a very useful chart to assess if you need a CMS at all, on which you can score the weight of your current content management problems.

Highly recommended. You could read this in a day or two and feel confident in embarking on a content management project.

Claire Pillar
Web Coordinator
North Cumbria NHS

The information society: a study of continuity and change. 4th ed
John Feather. London: Facet Publishing, 2005
ISBN 1-85604-497-1 240pp £22.95

This book was first published in 1994 and has reached its fourth edition. It is intended for students on courses in information studies, librarianship and communications. Within the broad area of communications, students of media, media production and multimedia will also find it useful. This book enables students to gain a basic understanding of the social, political and professional issues which are being generated by our growing dependence on computer networks in almost every aspect of our lives. To quote, "we have carefully avoided extended discussions of some of the underlying theoretical issues about the nature and concept of information." At one level the computer is an empowering tool which gives unprecedented access to information and to communications; at another it might be seen as an implement of control, giving unlimited power to those who control the information that it stores. Feather sees the central dilemma of the information society as how to resolve that issue.

This wide-ranging book covers the information-dependent society we live in by covering the historical, economic and political contexts. The historical context is that of writing, printing and other systems of communication. The economic context is the commodification of information and the means of its storage and transmission, covered here in chapters on the information market-place and access to information. Politically there is power to be derived from the possession of information and a loss of empowerment caused by its absence. Feather covers the political dimension in chapters on information rich and information poor, and information, the state and the citizen. The final chapter looks at the information profession: a domain delineated.

Feather considers the information paradox: that technology has made more information more available to more people than at any time, but made access to it more difficult. He goes into some depth in the latter part of the book about the debate on the boundaries of the legitimate information needs of a democratic state. The traditional book publishing model is covered in some detail because the process has become the paradigm of information transfer. Feather tracks the critical shift in structure of the information market place; that the commodification is extended from the product to its contents and the means of accessing its contents.

Naturally the Internet is covered, and he contrasts the power of the internet as against the power of transnational media groups: "The ubiquitous URL is the symbol of the revolution."

He is positive, as the potential for greater access to information has been created by the very technologies which have made it possible and perhaps even necessary to restrict that same access.

The most interesting section for me was that on data protection and personal privacy, reflecting on the use and holding of personal data by the public and private sector. He does not mention the amount of information held on 'loyalty' cards, perhaps of concern when £1 in every £8 spent on shopping is spent at Tesco supermarkets. He is correct in stating that intellectual property, data protection and freedom of information all exemplify different aspects of the three-way relationship between information itself, that state and the individual citizen. The state's role here has been to intervene in the chain of communication between the source and the user of information.

The final section on the information professional notes that the dissaggregation and recombination of skills…is typical of the changes brought by computers in the information world; traditional boundaries are being eroded and there is convergence between traditionally separate activities.

I hope this book will have further editions. There are so many developments of relevance. Perhaps the next edition will find newspapers declining even further; in this edition Feather talks about media finding ways to survive as newspapers in a society in which they are not the preferred news medium. Personally I hope ID cards do not get adopted here, but if they do get through, Feather will no doubt have a useful perspective on them. There will also be in place by then the electronic patient record for the health service (assuming this will be one of the public sector IT projects that succeeds). This includes the concept of an electronic "sealed envelope" into which will be placed any information that patients want to be kept confidential.

Feather sees the critical test of freedom of information in the UK as "whether the ingrained culture of secrecy in British government will change; early indications are not uniformly encouraging. He mentions that a new argument is that the control of flows of information is an essential weapon in the 'war' [sic] against international terrorism and criminality. In the scholarly publishing world there is the development of open access journals.

Rather than a bibliography of endless sources, there is a useful note on further reading. This guides the novice to the best books on the various topics discussed as well as what to read regularly.

As with all Facet books, the production is of a very high standard; well typeset with elegant fonts, and sturdy binding even in paperback. The index is decent, although it misses knowledge management, which is discussed briefly and some readers may wonder about

- are we in a knowledge economy or an information society?

Facet Publishing's website includes a sample chapter at time of writing: Chapter 5: information rich and information poor [http://www.facetpublishing.co.uk/286.pdf]

Claire Pillar
Web Coordinator
North Cumbria NHS

The Academic Library 2nd ed.
Peter Brophy. London: Facet Publishing, 2005
ISBN 1-85604-527-7 248pp

This introductory textbook is described as providing "a comprehensive overview of the changing functions of higher education libraries and the organisational cultures in which they operate". Topics include the historical context, the library within the institution; the increasingly heterogeneous user population; the impact of ICT, library systems and networks; resource and access management issues; performance indicators and the measurement of impact; the academic library building; human resources and professional matters; and future trends.

The approach taken for much of the book is primarily historical and focused on the UK. For example chapter 5 "The impact and opportunities of ICT" is devoted to a summary of the many government funded e-lib projects of the nineties, many of which are listed by name. However I looked in vain in the index for a reference to open url, SFX, deep linking or link resolver or indeed metasearch, cross searching or single search. However these techniques are widely used in academic libraries today to provide direct access to content in multiple sources. This functionality is also a key selling point for major library management systems yet no reference to it is made in chapter 8 "Library systems".

Also not to be found in the index was any mention of Google or indeed search engines as a concept. However any discussion of undergraduates or information skills must be incomplete without recognition of the impact widespread access to the internet and to search tools must have.

After so much history I turned with particular interest to the last chapter "The academic library of the future". Brophy identifies headline trends in the context of the opportunities and threats they offer. The section on information resources discusses the future of electronic v. print, the development of eprint repositories and their impact on the publishing industry, the role of electronic books for textbooks and scholarly monographs and the concept of learning object repositories. Then he touches on the future impact of household gadgetry, televisions and games monitors

(with a passing mention of the electronic fridge) and mobile communication devices as delivery mechanisms. In terms of the users of the academic library Brophy points to the increase in higher education with the prospect of a fifty percent graduate population.

The concept of lifelong learning and the continual need to update knowledge and skills must bring with it a requirement for information services. Virtual learning environments are meeting some of these needs. However the expensive failure of the e–university project and the recognition that most learners do not like entirely online learning indicate that the model for providing mass support for learning in the future has not been as yet developed and certainly the role of the academic library as we know it in this is not yet known. However in case the reader is disheartened by this uncertainty, Brophy concludes by reminding us that libraries are services with the aim of enhancing learning. He states that it seems "inconceivable that centres of information handling expertise and guides to ever-burgeoning information resources of the world will not be needed". Let us hope this is so.

Helen Edwards
London Business School Library

Books waiting for review

The following books from Facet Publishing are waiting for reviewers - please contact Peter Chapman if you would like to do any of them.

Pedley (ed) **Managing digital rights a practioner's guide**

Melling (ed) **Supporting e-learning a guide for library and information managers**

Press Releases

Emerald 2004 Statistics Show Usage Continues to Climb

Emerald Group Publishing, Ltd. has released its 2004 usage statistics, showing a seventeen percent increase in usage by customers worldwide. On average over 600,000 articles were downloaded each month, affirming market demand for Emerald content in all areas of management information.

Not surprisingly, the most popular spot on the Emerald Website is a weekly sample of two free journals. Widely known as Journals of the Week, this feature offers existing and potential customers full access to a rotating range of titles from the Emerald portfolio. The journal that experienced the highest number of article downloads in 2004 was the European Journal of Marketing.

For fourth consecutive year, the top-ranking article in the Emerald database was "From Marketing Mix to Relationship Marketing: Towards a Paradigm Shift in Marketing" by Christian Grönroos. It was published in Management Decision, which ranked number one among all journals in 2003.*

"Usage statistics have emerged as an industry standard for companies who wish to understand their customers," said Gill Crawford, Head of Corporate Communications for Emerald. "Additionally, librarians and information professionals will find these statistics a critical tool for evaluating and planning their service subject coverage."

Usage statistics are also valuable indicators for journal editors and members of the research community. Authors wishing to write articles for 2005 publication will benefit from focusing their efforts on management, marketing, and motivation, as well as knowledge management and business intelligence topics. Researchers looking for areas to investigate will be well served to consider these hot topic areas to ensure their findings yield useful results with immediate insights.

Reflecting Emerald's international customer base, top usage came in from Europe, Malaysia, and Australia, lending a global perspective to the statistics and subject interest areas.

*The top download Grönroos article can be viewed at http://images.emeraldinsight.com/emerald/pdfs/gronroos.pdf until 22 April 2005.

McGraw-Hill and Xrefer Establish a Partnership for Online Reference

Xrefer, provider of online ready reference services to libraries worldwide, has announced that leading reference publisher McGraw-Hill has agreed to license some of its key reference titles for inclusion in the Xreferplus service. This brings the number of publisher partners currently working with Xrefer to 42.

Xreferplus is a completely customizable ready-reference solution for corporate, academic and public libraries. It features full-text, aggregated content from hundreds of reference books, with 2 million entries covering such subjects as medicine, art, technology, business and more. By combining content from multiple publishers with unique cross-referencing technology, Xreferplus delivers authoritative answers fast, wherever and whenever library users need them.

Free, 30-day trials are available at http://www.xrefer.com/freetrial.

This new agreement will initially see McGraw-Hill licensing some of its bestselling reference titles for integration into the Xreferplus service, including some titles which have never been available in digital form before. Science, Technology and Architecture titles will be available in full text, complete with line drawings and illustrations, biographical entries and conversion tables.

Chad Davine, Vice President of Content and Market Development at Xrefer said of this development:

"We believe that McGraw-Hill is an example of academic reference publishing at its best, and we are delighted to be working with them in this way. It's a win-win situation: by partnering with us, McGraw-Hill will ensure that their works will be seen by thousands of people at institutions around the world, and by integrating their content into Xreferplus we will be enhancing our service and allowing our patrons access to these great titles."

Xrefer and A&C Black team up for Who's Who Online

xrefer and A&C Black are delighted to announce that Who's Who 2005 and the full Who's Who archive Who Was Who 1897-2004 are now available to search and browse on the xreferplus Specialist Reference platform.

Who's Who is the essential annual resource for accurate, detailed and up-to-date information on people of influence and interest. It was first published in 1849 and contains over 30,000 biographies of people in all walks of life. New names for 2005 range from famous faces like tennis player Tim Henman and movie star Nicole Kidman to Thomas Barnicoat, Chief Executive of TV production company Endemol. Each entry is carefully updated annually.

Institutions whose libraries purchase access to Who's Who 2005 online via xrefer will enable their users to search and browse for anyone whose biography has been included in the latest edition of the famous reference work. In addition, users will enjoy full access to the 10 volume archive of Who Was Who 1897-2004, which is fully integrated with the latest edition of Who's Who. Who Was Who 1897-2004 contains details of over 100,000 deceased biographees and dates back over 100 years to 1897 and is no longer available to purchase in print in full.

With the addition of xrefer's unique browsing interface, users can browse for biographees according to their profession, company or institution, education, address, birthday, birth year, recreations or decorations. They can also see which entrants are new in the latest edition of Who's Who. More information is available on the xrefer website at http://www.xrefer.com/ww2005.

Jill Coleman, Managing Director at A&C Black said of this partnership:

"Who's Who online is a fantastic package - a highly respected, accurate and up-to-date biographical resource on those of influence in our society, accessed via the versatile and user-friendly xrefer interface. This is a unique and invaluable resource and an essential purchase for libraries and companies."

John G. Dove, CEO of xrefer added:

"Who's Who is a well established and venerated giant in the world of reference publishing. We are delighted to be partnering with A&C Black to provide new browsing and searching functionality for Who's Who, and to bring it to an online audience for institutions."