



Internet applications, sites, trends and happenings

David Raitt

draittd@estec.esa.nl

This column aims to draw your attention to various interesting Web sites which I have come across and which might appeal to you, and to keep you up-to-date with news and views on Internet trends, developments and statistics. It offers essentially a personal selection rather than comprehensive coverage.

24 Hour museum newsfeed

Britain's National Virtual Museum, the 24 Hour Museum (<http://www.24hourmuseum.org.uk>), has developed a free newsfeed service to spread the word about UK museums, galleries and heritage attractions even further round the World-Wide Web. Using XML technology widely used in the Web sector, the new 24 Hour Museum service offers at least two museum and gallery stories every day. The stories come straight from the pages of the Web site and they are available for partner Web sites to link to.

Newsfeeds enable Web sites to reach out further with existing content to new audiences. The material also could have wider use, for instance with tourist boards, arts bodies and even the wider cultural and higher education sector. Members of the public can also read the stories directly on their own computers using free 'feedreader' programs freely available on the Net. The newsfeed is essentially an output of content from a Web site in a simple Web language, in this case RSS 2.0, that other users can link to for themselves. Newsfeeds are sometimes used by Web sites lacking resources to write their own content. In many cases, these feeds are given free, because the originators of the feed have a marketing or strategic aim it is important to publicize.

The 24 Hour Museum, launched in 1999, is an independent charity funded by the Department of Culture, Media and Sport (DCMS) through Resource, the Council for Museums, Archives and Libraries. It aims to promote UK museums, galleries and heritage attractions and encourage people to actually visit them around the country and see exciting activities all over the UK, all year round.

The Web site's growing database includes over 2800 entries and visitors to the site can search for what's on in the UK by place, date or by subject. The site features regular news and exhibition stories on the site, along with Internet trails produced in partnership with museums and galleries.

The new service has been available in pilot form on the 24 Hour Museum for several weeks, but a structured evaluation of the mechanics of how the feed works was carried out prior to publicizing the new facility. The feed has been offered in the first instance to <http://www.newsnow.co.uk>, a 'news agglomeration' site, to see what happens to 24 Hour

Museum site usage figures and from just this one user there has been a rise in new users to the site.

Just give me the facts, ma'am

Hooked on Facts.com was founded by Arthur W. Della Ratta in June 2003. Inspired by his young niece's reaction to a quick random-loading fact code that he had developed to keep her entertained, Della Ratta believed others would enjoy it as well. Hooked on Facts was developed to provide Internet users with quick, easy and enjoyable entertainment from some of the most amazing and unique facts ever uncovered. Claimed as the Internet's first dedicated fact engine, Hooked on Facts generates thousands of random interesting, as well as useless, facts. Users can add new facts of their own and can also receive each week a newsletter stuffed with some of the most amazing facts ever uncovered. Check it out for yourself at <http://www.hookedonfacts.com>.

New kids on the Net

Nielsen//NetRatings, the global standard for Internet audience measurement and analysis, reports that more than 27 million Internet users between the ages of 2 and 17 logged online from home in September 2003. Twelve million children aged 2 to 11 in the US accessed the Internet from home while 14.9 million teens aged 12 to 17 connected online. In September, kids between the ages of 2 and 17 represented 21% of active at-home Internet users or one out of five Web surfers. Mirroring the offline world, the 2 to 17 and 12 to 17 age demographics have markedly different on-line tastes. Most popular with the 2 to 11 age group was Diva Starz, a Mattel company Web site for the Diva Starz line of dolls. Of the site's audience, 54% consisted of kids aged 2 to 11. Second was ToonTown Online, a Disney company site, with 47% of the site's at-home audience. Polly Pocket, also a Mattel company site, picked up 46% of its audience, the third highest percentage, from this age group. Among teens, Originalicons.com, a popular destination for downloading instant messaging icons, had the highest concentration of those aged 12 to 17, making up 78% of its audience. Blunt Truth, an on-line game site, had 76% of its audience from the 12 to 17 age group, while Teen People had 74% of its audience coming from the 12 to 17 age demographic.

Nielsen//NetRatings also highlighted the US regions with the greatest concentration of Internet savvy kids. Leading the way is Salt Lake City, which had 255000 Internet savvy kids, representing 25% of the region's at-home Internet audience. In Cincinnati, 24% of the city's at-home Internet users, or 216000 were children aged 2 to 17. Boston filled the third spot with 633000 children going on-line from home. Rounding out the top five were Sacramento and Phoenix with both markets garnering 22% of the total at home audience. Read more at http://www.nielsen-netratings.com/pr/pr_031021.pdf.

Move towards higher screen resolutions

Internet users worldwide are moving towards higher screen resolutions according to WebSideStory's StatMarket and this trend is enabling Web developers to build more effective sites. As many Web developers know, 800 x 600 has been the dominant screen resolution worldwide for many years with 66% of users in January 2000, compared to 42% today. It was finally surpassed for the first time in April of this year by the larger 1024 x 768 resolution.

Today nearly 60% of Web users have a screen resolution of 1024 x 768 or higher, compared to only 34% in January 2000. This may be indicative of larger monitors and the trend is important because it enables Web developers to build more effective sites, with more content on the page, according to Geoff Johnston, vice president of product marketing for StatMarket, a leading source of data on global Internet user trends. 'For many years, developers and content providers have wished they could put more on the page, but were

constrained by smaller screen resolutions,' Johnston said. 'That is beginning to change and it's interesting to note that U.S. Web users lag behind the rest of the world in this area.'

Read more at

http://www.statmarket.com/cgi-bin/sm.cgi?sm&feature&week_stat

Who owns whom?

Yahoo! recently announced the completion of its acquisition of Overture. As a result of the merger, Overture is now a wholly owned subsidiary of Yahoo!. Overture is the leading provider of commercial search with more than 100000 active, paying advertisers globally. Overture's assets also include an extensive affiliate distribution network that both Yahoo! and Overture are dedicated to maintaining and enhancing. In addition, Overture has built a world-class technology infrastructure in commercial and Web search, as well as an extensive intellectual property portfolio. It also operates the AltaVista.com and AlltheWeb.com Web sites.

'We are excited to combine the two companies to build the largest position in the rapidly growing Internet advertising market,' said Terry Semel, chairman and chief executive officer, Yahoo! Inc. 'We believe the combined companies will provide the most diverse set of integrated marketing solutions on the Web for small and large businesses, as well as extend our advertising capabilities across the Yahoo! network and around the globe.'

The combination also furthers Yahoo!'s objective of becoming the leading end-to-end integrated search provider, combining assets capable of generating, distributing and monetizing search results. The two companies anticipate that they will be able to expand marketing opportunities on the Yahoo! network through the expansion of commercial search into vertical properties such as shopping, travel and yellow pages and integrating contextual advertising throughout Yahoo!'s network. In addition, Yahoo! will be able to leverage Overture's efficient, scalable marketplace by offering Overture's large base of advertisers additional Yahoo! services, and by expanding their combined products internationally.

So now AltaVista, AlltheWeb, Inktomi and Overture are all owned by Yahoo! While the advertising market is what pays for the search engines, the real question is which of these search engines will continue and at what sites? For now, AltaVista and AlltheWeb continue to be available at their historic locations, and they may share the same underlying database very soon. Already, AlltheWeb has lost a few search features, such as the URL Investigator which no longer displays the number for the links, but overall both still work with all of their old search features. Inktomi still is the back-end search engine at MSN Search and remains available at HotBot.

Full details about the merger can be found in the official press release at

http://www.corporate-ir.net/ireye/ir_site.zhtml?ticker=OVER&script=410&layout=0&item_id=456317

Disappearing act: Now you see me now you don't

As we all know, Web sites come and go. Greg Notess, writing on his Search Engine Showdown site (<http://www.searchengineshowdown.com/newsarchive/000724.shtml>), says that he has just noticed that Intelliseek's InvisibleWeb.com directory of hidden or invisible Web databases and resources now just redirects to Intelliseek's ProFusion meta search engine (<http://www.profusion.com/>). He says that it looks like they may have made the change on 13 August or earlier. While the directory had not been updated for a while, it was useful as a place to look for some hidden Web resources, especially databases. Some of those databases are still searchable within ProFusion's search groups, but it was the directory function that was most valuable. For a replacement, try Price and Sherman's

Invisible-Web.net (<http://www.invisible-web.net/>) or CompletePlanet (<http://www.completeplanet.com/>) instead.

I have also just noticed something similar. In preparing for this column I often checked the NUA.com site (also nua.ie), which claimed to be world's leading resource for Internet trends and statistics. Well, I noticed that I was redirected to the CyberAtlas site at <http://cyberatlas.internet.com/>. So presumably this means that CyberAtlas has taken over NUA recently, though I didn't see anything obvious on the new site to this effect. Since its inception in 1996, CyberAtlas was created in 1996 and was acquired by internet.com in August 1998.

CyberAtlas is billed as the Web marketer's guide to on-line facts. Do you need to know how many people use high-speed Internet connections or the number of on-line users in any one of more than 50 nations or what percentage of the Web is devoted to porn sites (it's 12%!)? The site provides readers with valuable statistics and Web marketing information, enabling them to understand their business environment and make more informed business decisions. CyberAtlas gathers online research from the best data resources to provide a complete review of the latest surveys and technologies available.

For quick statistics, check the CyberAtlas Stats Toolbox. The toolbox is divided into sections on Internet traffic, advertising statistics, demographics, e-commerce, statistics about Internet access and surfing habits, and weekly and monthly data. The geographic section provides data on how many people are on-line, either worldwide or in your own corner of the globe. For stuff on the cutting edge of Internet technology, have look in the hardware and wireless sections.

Lies, damn lies and statistics

Talking of statistics, take a look at NationMaster.com (<http://www.nationmaster.com>), a massive central data source and a handy way to graphically compare nations. Using a special form the user can generate maps and graphs with ease on all kinds of statistics, selecting only those countries you want to include. You can also view profiles of individual countries including their maps and flags, as well as a full encyclopaedia.

Currently some 650 statistics are provided and this number is increasing all the time. You can see how many murders, rapes or burglaries take place where; or population figures; or marriage rates; or drinking water availability; health stats; airports with paved runways – pretty much anything in fact. The most murderous nation is the USA (closely followed by South Africa), many more rapes take place in the USA than anywhere else (though South Africa is catching up). In fact the US ranks number 1 (i.e. the worst) in quite a few crime areas – most people prosecuted, most assaults, most car thefts, most females prosecuted and so on). South Africa is the most trigger happy with more murders there with firearms than anywhere else. Bangladesh and Nigeria are the most corrupt nations. Norway, Finland and Australia are the best educated.

Features include better visualization with scatterplots showing the relationship between any two statistics you care to name; pie charts showing proportions; printable versions of all graphs; and colour-coded maps with legends explaining what the colours mean. Clicking on the map gets you facts and figures for that country and the encyclopaedia with 150000 entries gives you much additional information. NationMaster also lets you look back in history with the GDP per capita statistics for the years 1820, 1900, 1950 and 1973. You can also look ahead to 2050 for projected population growth rates.

All in all this is a rather interesting site to while away some time – and it will help you with trivial pursuits! You can even see who is using the site. Would you guess Lithuania is using

it the most at the moment of writing – with Bermuda in second place? South Africa ranks down in 46th place!

Getting rid of spam (or trying to)

I've included items before on the problems and solutions to spam and now a new study has found that the waves of spam flooding into inboxes are beginning to turn people off e-mail (<http://news.bbc.co.uk/1/hi/technology/3209189.stm>). The study, by US think tank Pew Internet, says that some 25% of people said they were using e-mail less because they were receiving so much junk.

More than half of all e-mails are estimated to be junk with the percentage of junk messages growing at a phenomenal rate over the past 12 months. A year ago, spam accounted for just 2.3% of all e-mails, according to experts. By May this year, the figure was an astounding 55%. But, unfortunately perhaps, there were signs that spam works, with enough Americans taking up offers in unsolicited e-mails to justify the cost of sending thousands of messages.

Figures show that:

- 52% trusted e-mail less
- 76% were bothered by obscene or offensive content
- 75% were bothered that they cannot stop the flow of spam
- 33% have clicked on a link in an unsolicited e-mail
- 7% have ordered product from spam mail

You can read or download the complete report at <http://www.pewinternet.org/reports/toc.asp?Report=102>.

However, two developments in the US might help the rest of us. First, a California court has fined a marketing firm \$2m for sending out unsolicited e-mails telling people how to spam and now the US Senate has just voted (97 to 0) to outlaw deceptive 'spam' e-mail, and set up a 'do-not-spam' registry for those who do not want to receive unsolicited commercial e-mail. Internet 'spammers' who flood e-mail inboxes with pornography and get-rich-quick schemes could face jail time and million-dollar fines under the bill.

The vote marks the first time the Senate has taken action against an online scourge that now accounts for 50% of all e-mail traffic, frustrating consumers and costing businesses bandwidth and productivity. The bill would not outlaw all unsolicited commercial e-mail, focusing instead on the fraudulent or deceptive messages estimated to make up two-thirds of all unsolicited commercial e-mail.

Marketers who falsify return addresses or routing information, hide their pitches behind misleading subject lines such as 'Re: your request' or promote body-enhancement pills or other fraudulent products would face jail sentences of up to a year and fines of up to \$1 million; repeat offenders could face jail terms of up to five years. Marketers would also have to label sexually explicit messages to allow users to filter them out.

The bill would also prohibit marketers from sending unsolicited messages to consumers who place their e-mail addresses on a 'do-not-spam' registry, similar to the popular 'do-not-call' anti-telemarketing measure launched earlier this month by the Federal Trade Commission. Other common spammer tactics, such as hijacking users' identities, using multiple accounts to evade filters and sending messages to millions of randomly generated e-mail addresses, would be outlawed as well.

Although more than half of US states have passed anti-spam bills of their own, many of

which set tougher regulations for marketers, the bill would pre-empt most state laws, but would allow states to set higher penalties for deceptive or fraudulent activity if they wished. Read more at <http://www.cnn.com/2003/ALLPOLITICS/10/22/congress.spam.reut/index.html>.

It also seems that better filtering would help. Paul Graham uses a technique based on Bayesian filtering which uses probability to work out if a mail is junk or real. The system is based around a huge corpus of junk and spam mails that Graham gathered over a few months. These thousands of messages have been statistically analysed to extract the top 15 features that define them as spam.

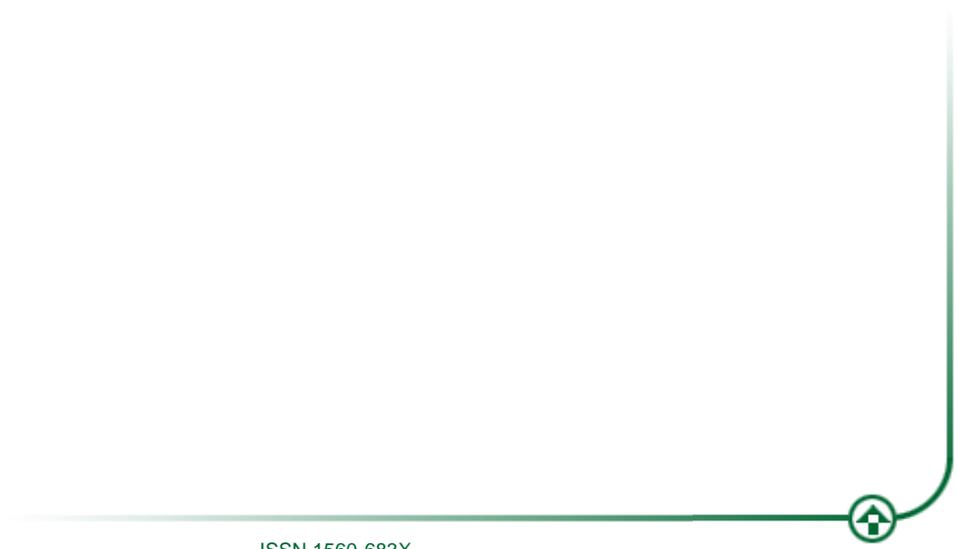
Any incoming mail is scanned to see how many of these defining characteristics it possesses. The list of defining features includes some words, such as 'teens', but others were less obvious and include formatting codes and routing information found in e-mail headers. Bayesian filters, such as CRM114 (<http://crm114.sourceforge.net/>), are virtually 100% accurate in spotting spam. Read more at <http://news.bbc.co.uk/1/hi/technology/3014029.stm>.

About the author

Dr David Raitt is senior technology transfer officer with the European Space Agency in the Netherlands. His work involves finding applications for space technologies in non-space sectors, particularly those useful for improving everyday life. An information scientist by education and training, David is also editor of *The Electronic Library* and chairman of the Internet Librarian International conferences.

Disclaimer

Articles published in SAJIM are the opinions of the authors and do not necessarily reflect the opinion of the Editor, Board, Publisher, Webmaster or the Rand Afrikaans University. The user hereby waives any claim he/she/they may have or acquire against the publisher, its suppliers, licensees and sub licensees and indemnifies all said persons from any claims, lawsuits, proceedings, costs, special, incidental, consequential or indirect damages, including damages for loss of profits, loss of business or downtime arising out of or relating to the user's use of the Website.



ISSN 1560-683X

Published by [InterWord Communications](#) for Department of Information Studies,
Rand Afrikaans University