



Internet applications, sites, trends and happenings

David Raitt

drait@estec.esa.nl

This column aims to draw your attention to various interesting Web sites that I have come across and that may 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.

Wikipedia – open encyclopaedia

Wikipedia (<http://www.wikipedia.org>) is a multilingual project that aims to create a complete and accurate open content encyclopaedia that is free of charge. Started in January 2001, the encyclopaedia already has over 100000 articles in the English version and more than 37000 in other languages including Danish, Dutch, Esperanto, German, French, Italian, Polish, Portuguese, Spanish and Swedish.

Wikipedia is a public WikiWikiWeb, a Web site where anyone can edit any article at any time without having to register as a user. Users not only contribute entries, but also build on each other's edits, and vandalized articles are quickly repaired by restoring an older version of an entry. In Wikipedia's second year, thousands of volunteer editors from around the world added 80000 entries to the English version and 33000 entries were made to the other language editions of Wikipedia. With entries being made 24 hours a day, seven days a week, this surge in growth has made Wikipedia the largest and fastest growing open content encyclopaedia in the world and the largest WikiWikiWeb. Wikipedia content is available under the terms of the GNU free documentation license, which ensures that anyone may re-use the entries on the site in any way they wish, including commercially, as long as they preserve this right in their own versions as well and credit Wikipedia as the source.

Internet entrepreneur Jimmy Wales and philosopher Larry Sanger founded Wikipedia. The search engine of Wales, namely Bomis.com has supplied the financial backing and other support for the project, and Sanger led the project during its first year as a full-time paid editor of Wikipedia. They both attribute Wikipedia's success to the presence of a strong core group of well-educated and articulate contributors from around the world who together maintain community standards of quality and neutrality. From the beginning, Wales and Sanger believed that it was absolutely necessary that all participants be committed to what they call 'neutral point of view' rather than taking stands on issues of controversy. Participants work together to prepare descriptions of a controversy that are fair to all points of view.

In December last year, the Wikipedia team also launched a sister project called Wiktionary (<http://www.wiktionary.org>), a free multilingual dictionary and thesaurus.

Internet Public Library

The Internet Public Library (IPL) is an all-digital public library that provides library resources and services to over 1,5 million Internet surfers from around the world every month. IPL (<http://www.ipl.org>) is the first public library of and for the Internet community. It is basically an experiment that tries to discover and promote the most effective roles and contributions of librarians to the Internet and vice versa. The IPL began life at the School of Information and Library Studies at the University of Michigan in the 1995. The idea was twofold: first to ask some interesting and important questions about the interconnections of libraries, librarians and librarianship within a distributed network environment; and second to learn more about these issues by actually designing and building something called the Internet Public Library.

Although IPL began as a student project, it is now an established entity that includes a small full-time paid staff. At the same time, there are always students and other people working on special and ongoing projects for IPL. The objective is not to replace libraries, but to offer an alternative. In 2002, a new design for the Web site was unveiled. The new site was created to provide IPL visitors with easier navigation, enhanced searching capabilities and updated graphics. New features being offered include searching tools that have a 'search this site' function, navigational menus that include a top 'you are here' menu and a sidebar menu, and a consistent layout and graphical design that extends over the whole site, including the youth areas (KidSpace and TeenSpace). IPL visitors can continue to engage in such activities as browsing on-line newspapers, magazines or books, finding authoritative Web sites on a particular subject or using the 'ask a question' service to get help with their research. Some individual resources have been renamed, making them more accessible to visitors.

Librarians Index to the Internet (LII)

The Librarians Index to the Internet (<http://lii.org>) is a searchable, annotated subject directory of more than 11000 Internet resources selected and evaluated by librarians for their usefulness to users of public libraries. Both librarians and the general public use the index as a reliable and efficient guide to Internet resources.

The LII began in 1990 as a reference librarian's Gopher bookmark file. In 1993, it migrated to the Berkeley Public Library's Web server as the Berkeley Public Library Index to the Internet. In late 1996, a search engine was added together with Library of Congress subject headings to make lii.org compatible with major standards such as Z39.50 and MARC. A virtual workspace to facilitate team-based record creation and maintenance was developed. In March 1997, the Berkeley Public Library Index to the Internet was moved to the Berkeley SunSITE and was renamed the Librarians Index to the Internet (lii.org). Since October 2000, the Library of California has provided much of the operational funding, while hosting is provided by the UC Berkeley SunSITE.

Scientific images on-line

The SERIMEDIS scientific image server (<http://www.serimedis.tm.fr/anglais/index.htm>) is an interface to the video and picture libraries of major French research institutes. The service offers thousands of videotapes and images available from a searchable catalogue using multi-criteria and with the help of a hypertext index. Although the service is open to all, only subscribers have access to detailed information and can order products. However, subscription is free of charge.

SERIMEDIS contains content from the following institutions:

The INSERM Image Bank offers photographs, documents and videos made in the 270 research units of the *Institut National de la Santé et de la Recherche Médicale* (National Institute for Health and Medical Research). Included are 14000 photographs covering the main research fields relevant to current medical and scientific issues and 2000 filmed microscope image sequences on molecular biology. Some 460 researcher-photographers are constantly adding to the collection.

The *Institut Pasteur's* photo library contains more than 8000 pictures tracing the history of the *Institut Pasteur* and its founder Louis Pasteur. It covers the full range of the institute's modern-day research in microbiology, immunology and developmental biology. The collection also includes pictures from associate Pasteur institutes outside mainland France, showing the pathologies of their countries of origin.

The Indigo Base from the *Institut Français de Recherche Scientifique pour le Développement en Coopération* (French Institute of Scientific Research for Development in Cooperation) is a picture library of the tropical regions, covering Africa, Latin America, Asia and Pacific. It contains 6000 pictures on a wide range of subjects, among which are agriculture, environment, ethnology, flora, geography, geology, hydrology, nutrition, oceanography, plant genetics, researchers in the field, seafloors, sociology, soil science, street art, tropical diseases, urbanization and so on.

Assistance Publique-Hôpitaux de Paris (the Paris Health and Hospitals Authority) has a photo and video library containing over 250000 photos and 400 videos. The main subjects covered are medical and surgical healthcare techniques, the patient-medical staff relationships, hospital architecture, patient reception and accommodation structures, laboratories, ambulances and activities in the ward and the medical professions. One large collection of photographs traces the history of the institution from 1900 to the present day.

The *Musée d'Histoire de la Médecine* (Museum of Medical History) in Paris was established in the 18th century and has been open to the public since 1954. It has a rich and varied set of collections, including surgical instruments representing the surgeon's art from the 17th to the 20th century. Other collections include the doctor's instrument kit, as well as engraving, lithographs, manuscripts, bookplates and photographs that illustrate the history of medicine.

Biologie sur Encyclopédie Électronique (Bee) is a terminological database that presents, defines and illustrates words, terms, expressions, abbreviations, symbols, concepts, etc. used in and relating to cellular and molecular biology, and their disciplines and areas of applications.

Internet scout

The Scout report is one of the Internet's longest-running weekly publications, offering a selection of new and newly discovered on-line resources of interest to researchers, educators, and anyone else with an interest in high-quality on-line material. The report, first published in 1994, is available both on the Web site (<http://scout.wisc.edu/>) and in e-mail form via mailing list subscriptions. Past issues of the Scout report, as well as past issues of the discontinued subject specific reports, are available. Every day, professional librarians, educators and content specialists filter hundreds of announcements looking for the most valuable and authoritative resources available on-line. Information about the best of what has been found is then summarized, organized and provided to the Internet community. The Internet Scout Project is located in the Department of Computer Sciences at the University of Wisconsin-Madison, and is sponsored by the National Science Foundation to provide timely information to the education community about valuable Internet resources. Daily and

weekly updates are offered for kindergarten children to graduate students to life-long learners.

In addition to the more general Scout report, three biweekly NSDL Scout reports provide readers with a selection of the newest and most interesting, high-quality on-line science, mathematics, engineering and technology resources. Available via free e-mail subscription or on the Web, the reports are part of the National Science Foundation's larger National Science Digital Library (NSDL) project. Areas covered by every report include the following:

- NSDL Scout report for life sciences: biology, zoology, ecology, botany and other life science topics
- NSDL Scout report for physical sciences: geology, chemistry, astronomy, physics and other physical science topics
- NSDL report for mathematics, engineering and technology: industrial engineering, calculus, algebra, geometry, civil engineering, applied mathematics, environmental engineering, computer sciences, human factors, hardware and software, and related topics. The reports cover information of interest to all ages of readers.

Select your own plot and characters

Whichbook.net gives readers an enjoyable and intuitive way to find books that match their mood. Instead of starting from the overwhelming choice of books available, whichbook.net starts with the reader and enables each individual to build the elements of that elusive 'good read' we are all looking for but do not quite know how to define. Whichbook.net offers the would-be reader choices that are not available anywhere else – plot types, main character types, the country in which the story is set, etc.

The standard way of organizing books for choice on shelves in a library, in a bookshop or on the Web starts from the available 'products', namely the authors, titles, publishers or genres. Using innovative software from Applied Psychology Research Ltd, whichbook.net enables, for the first time, a start at what the individual reader is looking for – particular subjects readers want to know more about but are unable to find or define exactly. Instead of just looking for a crime novel or a book by a favourite author, the reader can look for a book that is unpredictable, very romantic and a little bit sad, or a challenging book that is also funny with plenty of sex. You can choose books, for instance, that are happy or sad, funny or serious, safe or disturbing, expected or unpredictable, beautiful or disgusting, gentle or violent, conventional or unusual, optimistic or bleak. You can select the race, age, sexuality, gender, etc. of the characters. The plot can be full of conflict or success against the odds or full of twists and turns, and you can select the setting from little country maps. Whichbook.net finds titles that match your preferences – there are 20 million different individual permutations possible! The books in the database are all fiction and poetry in paperback, written in or translated into English and published since 1995. Books are read by one of a team of 150 people who are drawn from libraries and literature organizations and are specially trained to create the entries. To see a list of titles go to <http://www.openlibraries.net>.

Whichbook.net (previously known as Book Forager) was developed by Openlibraries Ltd. Openlibraries has been awarded £350000 from the New Opportunities Fund to expand whichbook.net over the next two years by increasing the number of titles on the system and by creating links between whichbook.net and public library catalogues on the Web. For more information, go to <http://www.openlibraries.net> or <http://www.whichbook.net>.

Gadget re-launched

South Africa's oldest on-line magazine devoted to hi-tech was formally re-launched at the end of 2002. *Gadget* magazine first appeared on the Web in February 1998, with a daily updated mix of reviews, opinions and insights into consumer technology. Through *Gadget*, consumers could continually stay informed about the latest market trends and know exactly what the new products are, what they do and whether they offer value for money. The magazine was originally published by Media Africa, which became part of the Acuity Group the following year. However, when founding editor Arthur Goldstuck departed from Acuity last year, he acquired *Gadget* back from the group. The magazine has been redesigned and is now published by his technology research organization, World Wide Worx (<http://www.theworx.biz/>). *Gadget* covers consumer technology, office hi-tech and general gadgetry that makes people's lives easier. Reviews get right down to the bottom line with the gadget five question user tests, enabling readers to see at a glance whether a product is ready to use, easy to use, functioning as it is supposed to, innovative and value for money. *Gadget* can be accessed at <http://www.gadget.co.za>.

What is that song?

Have you ever heard a song on the radio or TV and wondered what it was called and who was singing it? Well, a new service called Shazam identifies songs through your mobile phone anywhere in the UK. All you need to do is dial 2580 on your mobile phone and point it at the music source (e.g. the speakers) to tag the track being played. After 30 seconds, Shazam terminates the call and you receive a text message with the name of the track and the artist – sufficient information for a trivia quiz, to download the track from the Web or to enable you to buy the CD. Shazam has a database of some 1,6 million songs to identify a song and can do anywhere, even in noisy places like bars and cafés. It identifies an original track and any of its re-mixes. Shazam is partnering with UK-based mobile operators, radio stations, record labels and book sellers like Amazon.co.uk to make sure the service is as widely available as possible. They are also planning to take the service elsewhere. Find out more at <http://www.shazam.com> and take your mobile phone with you the next time you visit the UK!

Web news snippets

The Internet survey Web site <http://www.nua.com> often has some interesting facts and figures. Here are some of the most recent offerings.

Pop-up advertisements may be unpopular with Internet users but they are highly effective at generating response rates, reports ZDNet. According to a study conducted by Gartner, 78% of respondents claimed they found pop-up advertisements 'very annoying'. In contrast, only 49% of participants felt the same about banner ads. However, the survey indicated that the click-through rate for pop-up advertisements was almost twice that of banner advertisements. Gartner believes that the high response rate may be inflated because many Internet users are unsure how to shutdown pop-up windows when they appear. The research company predicts that once users learn how to close the windows, the response rates will decrease.

Spam accounted for over 40% of e-mail messages in the UK during December 2002, reports Silicon.com. According to the anti-spam firm Brightmail, the amount of spam in the UK during December was up 7% compared to November 2002. Adult mail accounted for 16% of all unsolicited e-mail in December. Financial services accounted for 26% of spam, while product promotions topped the list with 31%. Fraudulent spam, such as the infamous Nigerian (and Zimbabwean) e-mail letter, accounted for 6% of all unsolicited mail. InternetNews reports that spam costs corporate America close to \$9 billion in 2002, according to a new study from Ferris Research. Productivity losses accounted for as much

as 40% of overall financial losses associated with spam. Ferris estimates that office workers waste around 4,5 seconds deleting every e-mail message. Consumption of bandwidth and use of technical support time were also cited as costly efforts resulting from unsolicited e-mail messages. Ferris Research calculates that spam accounts for between 10–15% of all e-mail sent to companies. The research company estimates that spam is currently costing European companies around \$2,5 billion per annum.

Europemedia reports that 180 million e-mail messages were sent daily in Italy during 2002. According to research by the Italian ISP, Netscalibur, around 160 million messages were sent daily by business users, while home users sent 20 million on average. A total of 70 billion e-mail messages were sent in 2002, up from just 55 billion in 2001. Italian employees spent approximately two working hours a day sending and receiving an average of 70 messages a day.

New Books

Here are two new books from Idea Group Publishing (IGP).

Web Powered Databases, edited by David Taniar and Johanna Wenny Rahayu (2003, ISBN: 1-59140-035-X).

Database technologies have revolutionized the way organizations conduct business. They have simplified ordinary business processes and transformed commerce activities. Getting the most from these technologies is essential for every organization. The next generation of databases that are powered by the Internet are fast becoming an essential part of conducting business. Consequently, information technology researchers, teachers, students and practitioners involved in electronic commerce must have access to the most current research and practice surrounding Web-based databases. David Taniar and Johanna Wenny Rahayu's new book, *Web-Powered Databases*, offers a collection of the most up-to-date research findings and best practices surrounding Web-based database technologies. From providing practical, real-life applications to discussing the intricate theory necessary to be successful in designing and implementing these technologies, this book offers the most current information available.

As the Internet becomes more widely used as a source of information and as a mechanism of conducting business, Web-powered databases are becoming increasingly important. Since databases are basically computerized record-keeping systems (with the increased use and development of Internet technology), it makes sense to have a database system implemented on the Web to make the information more accessible. Web-powered databases have their own set of challenges and opportunities. Although performing data queries requires new applications do be done effectively on the Web, Web-mining (an application of data mining) offers real opportunities for organizations that use it. Web mining is basically a means for discovering patterns in user accesses on the Web. This information will be particularly useful in building a Web portal that is tailored for individual users.

This new book addresses the questions associated with this emerging technology. It specifically examines the kind of programming that should be used in the database design and data queries to answer the questions surrounding Web-powered databases. Additionally, because of the cost of these databases, organizations are interested in maximizing their investments by achieving optimal performance of their Web-based database.

Web-Powered Databases has 11 chapters, all authored by leading experts from eight different countries in the fields of Internet databases and data mining. It contains practical case studies and tutorials designed to help designers, students, teachers and researchers

understand the complexities associated with these emerging technologies. Additionally, it examines the theory behind successful query design, the architecture of a Web-based database and the important security considerations. This book is both a reference for the Internet generation's practitioners, researchers, students and teachers and a must read for anyone interested in the development of database applications.

David Taniar received his Ph.D. in computer science from Victoria University, Australia and is currently a senior lecturer at the School of Business Systems, Monash University, Australia. Johanna Wenny Rahayu received a Ph.D. in computer science from La Trobe University, Australia and is currently a senior lecturer at La Trobe University.

Building a Virtual Library, edited by Ardis Hanson and Bruce Lubotsky Levin (2003, ISBN: 1-59140-106-2).

The Internet has made information available on any topic, at any time, to anyone, and subsequently it has changed the nature and process of research. It has also made access to the most current information a priority for researchers and teachers. However, this increase in the available electronic and digital media has created difficulties for libraries as they attempt to create and maintain their digital collections. The decisions on what to include, how to organize available information and how to ensure access to the most current information are of paramount importance in academic libraries. Building a virtual library will provide a framework that helps librarians, researchers and administrators in the creation and maintenance of the new services and resources unique to the virtual library.

The development of a virtual collection is an expensive and time-consuming task. In the development process, virtual libraries face their own set of unique challenges. Virtual libraries are organic. Understanding the challenges of development is ongoing. These challenges range from content to interfaces, from digital video to geospatial infrastructures, from staffing to marketing. The unique set of challenges that an organization faces in creating and maintaining its virtual collection requires commitment from stakeholders at all levels. The function, role and operation of libraries have changed as they deal with electronic media. Libraries no longer own materials, they license them. However, new means of publication (such as electronic pre-print services and depositories of scholarly publications) promise to transform the methods by which scholars exchange and preserve the results of their work. This will, in turn, transform academic libraries.

From Web-based electronic catalogues to databases of electronic publications and media, the tasks involved in creating a virtual library are diverse and vast. Libraries have repositioned themselves in a virtual world, providing access to their repositories and services and undertaking massive and successful digitization efforts of text, images, sound and datasets.

The success of university libraries in the 21st century will largely depend on their ability to create efficient and effective virtual libraries. The role of the university librarian is changing with the changing technologies. The librarian of the 21st century must have technological savvy and be aware of all the information that is available. Building a virtual library presents both theoretical approaches and practical case studies within an academic environment. Librarians will have a greater understanding of how these new technologies change their environment and their roles. Academics and researchers will understand the importance of the virtual collection and administrators will benefit from the logistical staffing and organizational change discussions.

Ardis Hanson is the director of the research library at the Louis de la Parte Florida Mental Health Institute at the University of South Florida, USA. Bruce Lubotsky Levin is associate

professor and head of the Graduate Studies in Behavioral Health, the University of South Florida de la Parte Institute and the College of Public Health.

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 the Centre for Research in Web-based Applications,
Rand Afrikaans University