TECHFOCUS

All about search engines

EDWARD APURBA SINGHA

IMAGINE yourself inside a big library wherein you are frantically searching for a reference book in order to prepare your assignment. You could see it's like looking for a needle in a haystack. Amid this moment of desperation comes the library management software with the promise of help. This software enables you to trace to the exact location of and other information about your desired book.

Same is the case about a search engine which you need to browse the internet. The web has now become a virtual world of millions of websites. Your any attempt to sift any specific information from the web without a search engine is sure to end in vain. This is why search engines have become an integral part of the internet.

The generic name 'search engine' does not actually mean any electronic or mechanical device. It is in fact a specially designed web page that is able to collect information from other sites on the internet.

Based on keywords or a phrase, search engines accumulate information from different sources and keep an index of the words they find and from where they find them. Afterwards users are allowed to look for words or combination of words found in that index.

The immense popularity of modern search engines is because of the mammoth growth of websites. At the early stage, however, the case was not as it is today. Back then people had to find essential files through programs like Archie and Gopher. These programs kept indexes of files on servers which were connected to the internet.

Early search engines were capable of maintaining a relatively small index of a few hundred thousand pages. Documents and received inquires ranged between one or two thousand per day. Today, however, all top search engines will index hundreds of millions of pages and are capable of handling tens of millions of queries a day.

The most important and fascinating aspect of a search engine is its search techniques.

Some search engines look pretty simple. This humble look, however, hides the real complexity inside it. Special software robot called spider is the vital component of a search engine. Spider robots have to examine several pages in order to prepare a list of words -- a process called web crawling.

The spider first searches for the most popular pages and gleans information from them. This mechanism is slightly different in different search engines. The google spiders, for example, mainly identify two things: the location where they were found. They seek out words from the title, subtitles and Meta tags and prepare an index excluding the articles -- a, an,

wor the AltaVista wor k s in a different way by indexing every single word on a page. It does not overlook the articles and insignificant words.

Meta tags are responsible for accelerating the search

process. Spiders examine it for a brief introduction to the contents of a web page. In many circumstances it creates hassle for which robot

exclusion protocol was developed. This protocol is implemented in the Meta tag section at the beginning of a web page that guides the spider to leave the page.

Once spiders' action is accomplished, the search engine needs to store the information in a way that makes it useful. The search engine, which generates result comprising only words and its links, apparently weakens its effectiveness, because it is not possible to perfectly identify the exact use of words on a particular website or links associated with it.

In order to fix this problem most search engines store more than just the word and URL. An engine needs to determine the frequency of word emergence on a page. Then it assigns weight to each entry, with increasing values assigned to words as they appear near the top of the document, in sub-headings, links, Meta tags or in the title of the page.

Each commercial search engine adopts different strategy for assigning weight to the words in its index. This is one of the reasons why a search for the same word on different search engines will produce different lists, with the pages presented in different order.

As a user you can follow Boolean search technique in order to enrich your search experience. Some tips regarding this are given below:

AND It is used to join two or more words. For instance, type Bangladesh + Sports and press enter. The search result will come up with information on Bangladeshi sports. Some search engines substitute the operator "+" for the word AND.

OR It enables search engines to show result that encompasses one of the terms joined by OR.

NOT It discards irrelevant topics from the search result. Some search engines substitute the operator "-" for NOT.

FOLLOWED BY - One of the terms must be directly followed by the other.

QUOTATION MARKS Inside quotation marks words are treated as a phrase which produces more precise result. Boolean search creates a dilemma when a word has different meanings. In such a case, new search tactics such as concept-based search and natural language queries will add a huge advantage to make your search more interesting.

Text Search	Video Search	Image Search	Blog Search
www.google.com	video.google.com	image.google.com	Blog.google.com
www.yahoo.com	video.yahoo.com	www.sphere.com	www.sphere.com
www.alltheweb.com	www.youtube.com	www.picsearch.com	www.blogdigger.com
www.altavista.com	video.aol.com		www.bloglines.com
www.live.com	www.blinkx.com		www.blogdigger.com
www.ask.com	www.tubesurf.com		www.icerocket.com

PHOTOTECH

WORLD'S FIRST BIONIC HAND

Retired US Army Sergeant Juan Arredondo, a veteran of the war in Iraq, wears the world's first bionic hand with independently moving fingers on July 23 in New York City. Arredondo lost his hand following an IED explosion in Iraq and is now able to use the i-Limb which has finger "joints" that act like natural fingers.

PHOTO: AFP

TECHVIEWS

Lawsuit seeks to shut down Facebook

AP, Boston

THE owners of a rival social networking Web site are trying to shut down Facebook.com, charging in a federal lawsuit that Facebook's founder stole their ideas while they were students at Harvard.

The three founders of ConnectU say Mark Zuckerberg agreed to finish computer code for their site, but repeatedly stalled and eventually created Facebook using their ideas.

The lawsuit's allegations against Zuckerberg include fraud, copyright infringement and misappropriation of trade secrets. It asks the court to shutter Facebook and give control of the company and its assets to ConnectU's founders.

Facebook has responded by asking a judge to dismiss the lawsuit. A hearing is scheduled for Wednesday in U.S. District Court in Boston.

Facebook started in 2004, a few months before ConnectU went online, and now has 31 million users, compared with about 70,000 users for ConnectU, based in Greenwich, Conn. Last year, Facebook turned down a $1 billion buyout offer from Yahoo Inc.

A spokeswoman for Palo Alto, Calif.-based Facebook declined to comment. But in court filings, Facebook's attorneys say ConnectU has no evidence for "broad-brush allegations" against Zuckerberg, and deny he pilfered his ideas for Facebook from his fellow Harvard students.

"Each of them had different interests and activities," they wrote. "Only one of them had

an idea significant enough to build a great company. That one person was Mark Zuckerberg."

Facebook and ConnectU connect college students and others online. Both allow users to post profiles with pictures, biographies and other personal information and create extended networks of people at their schools or jobs or with similar interests.
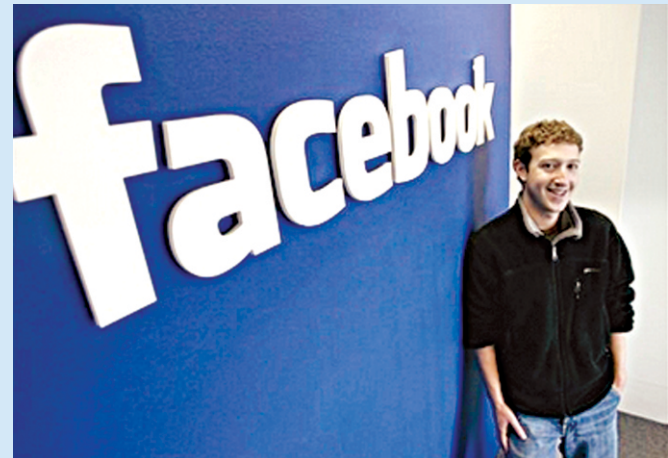
ConnectU originally filed suit in 2004, but it was dismissed on a technicality and immediately refiled. The lawsuit claims that

Facebook.com's mastermind, Mark Zuckerberg smiles at his office in Palo Alto, Calif.

in December 2002, ConnectU founders Cameron and Tyler Winklevoss who are brothers and Divya Narendra began to develop a social networking site for the Harvard community called Harvard Connection.

In November 2003, the three asked Zuckerberg to complete software and database work on the site. They repeatedly asked him to finish before they graduated in June 2004, and Zuckerberg assured them he was working hard to complete

it, the lawsuit says.

"Such statements were false and Zuckerberg never intended to provide the code and instead intended to breach his promise ... and intended to steal the idea for the Harvard Connection Web site, and in fact he did so," the suit alleges.

Zuckerberg launched Thefacebook.com in February 2004. ConnectU started its Web site in May of that year. By beating ConnectU to the market, Facebook gained a huge advantage, the lawsuit claims.

ConnectU's founders have written on their Web site about the "ups and downs" of their company history, including a programmer "who stole our ideas to create a competing site."

"But we've been troopers," they wrote. "At first we were devastated and climbed into a bottle of Jack Daniels for a bit, but eventually emerged with a bad headache and renewed optimism. We weren't going to lie down and get walked over like this."

TECHNEWS

Oracle Applications Day observed

SOFTWARE giant Oracle observed 'Oracle Applications Day' in Dhaka on July 25. Through this programme the company tried to publicise its upcoming software solutions for the next generation mission critical business activities.

In his inauguration speech, Sajjad Syed, business manager, MRD, Oracle Corporation, spoke about the achievements and challenges of the campaign, Oracle Applications Unlimited. He said the inside story of "our success is better information, leverage next generation technologies and customer-driven innovation".

In an informal conversation with StarTech, Sajjad hoped that Oracle would start its operations

in Bangladesh by the end of the year. "Oracle Corporation acquired 17 security certifications and in Bangladesh nearly all the renowned businesses including the GSM mobile phone operators currently use oracle technology", he added.

Irfan Parkar, Warid Telecom's deputy general manager IT, stated his experience as a corporate user of Oracle. He informed that the carrier has already successfully run Oracle's ERP solutions. "The success of any solution totally depends on the efficient management as well as selection of right consultants", he said.

Shantanu Mitra, AVP-sales, Metalogic Systems, told StarTech that they are planning to unveil its ERP solutions for the Bangladesh market. The

India-based company basically works on software transformation, insurance solutions and ERP solutions. Currently Metalogic has its presence in Japan and Europe. Following the speeches, two breakout sessions were organised in order to interact with Oracle customers. Conducted by Metalogic and IBCS-Primax, they included some important discussion sessions -- 'Oracle supply chain planning for profit optimisation', 'Oracle manufacturing for better resources utilisation', 'Oracle financials for enabling an insight driven enterprise' and 'Oracle's solution for human capital management'.

EDWARD APURBA SINGHA

Tech Jargon

IN today's digital era, it's quite easy to find yourself in situations where you're bombarded with hi-tech jargon that you've heard many times, but have no idea what they mean. The most annoying thing of all is the IT industry's love for acronyms!
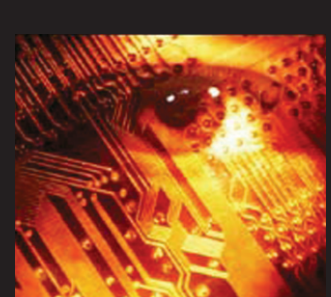
Keeping this in mind, during the next few weeks, we will feature a few of such technical jargon in a bid to come to the rescue of the poor layman's soul, so that the next time you hear someone say "my firewall doesn't let me access my ISP's server", you'll know what it means. So all you technophobes out there, watch out for this space every week!

Cyberspace: I can bet there's not a person out there who hasn't heard this word before. Loosely speaking, cyberspace is nothing other than the internet itself.

E-commerce: If you're a

businessman, you ought to know this one. E-commerce is the conduct of business over the internet, something that all big companies are beginning to gear towards.

ISP: Abbreviation for *Internet Service Provider*, it's

the nice old company that provides you the connection to the internet. May God bless them!

Network: A group of computers connected together to facilitate sharing of information and other

devices.

Server: No, it won't serve you biscuits and tea. A server is a central computer that "serves" other computers by helping them access files from the network, take printouts from a shared printer, and other such services.

WAP: Stands for *Wireless Application Protocol*. It's something that you'll find on the tip of every youngster's tongue these days. A service provided by most mobile phone companies, it is a system that allows you to access the internet through your mobile phone. That sounds exciting, doesn't it? However, in practice, the site to be accessed has to be specially designed to work for mobile phones, and very few sites have bothered to set up a WAP version!

NAHID AKHTER