

## Searching Google's Haystack

By Moshe Ofer

As an April Fools' prank, a New York university publicized a new degree it was offering: a master's in Google. Ray von Dran, dean of the Information School at Syracuse University, was quoted in the Library Journal as saying that the school would offer a program leading to a Master's Degree in Google Science (MIGS).

"Americans are Googling for hours every day, and they need leaders to show them how to make that activity more efficient," von Dran said (sic).

The joke was that since everyone was Googling now, it was time to offer expertise in the area, showing others how to search and actually find the thing they are looking for. Like all humor, there is a nugget of truth to the idea that you need a "master's in Google" to find what you seek on Google.

Google has an index of 10 billion pages. Finding a precise answer to a question posed to that much of a stack of information is akin to the "nee-



Moshe Ofer

*Moshe Ofer is president/CEO/director of Activepoint Ltd. His e-mail address is ofer@activepoint.com*

dle in a haystack" notion if you do find exactly what you want. It's just plain dumb luck.

Google is often the tool of students and perhaps business people with time on their hands. Without the help of ads that appear on the right side of the Google search page, a person must sift through the thousands of pages that are served up for a query.

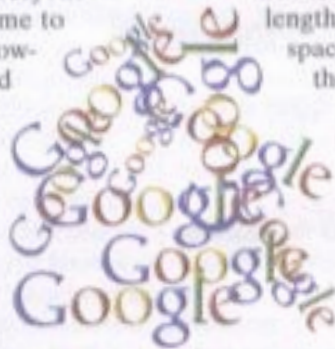
**What Google isn't.** Google is less of a system of finding what you want swiftly so you can have that information to make a business decision or buying decision. It is much more an "ad engine." If you pay the money, your information will be served up more prominently than those that do not. It doesn't mean the information being paid for is any more or less valuable.

This is akin to paying for shelf space in the retail food sector. Those that want their food product displayed at arm's length on the shelf pay more for space than those that end up on the bottom shelf.

The more pages of information that are served to the searcher, the better it apparently is for Google's revenue stream, which seems to be ad-based. According to the Internet Ad Revenue Report conducted by PricewaterhouseCoopers for the Interactive Advertising Bureau, Internet ad revenue for the first nine months of 2004 totaled slightly over \$7 billion compared with \$7.3 billion for all 2003.

One can assume that the harder it is for

**Continued on page 18**



Continued from page 17

## Searching Through Google's Haystack

someone to find something quickly from a search query, the more likely they will click on one of the growing number of sidebar ads that are listed. The ads are based on some of the keywords in the search string. Even these ads that are served may be irrelevant to what the person seeks, and they may be out of context.

The side banner ads are based on what is paid for, and may or may not solve what the searcher seeks. They may even lead the searcher in a different direction. So Google is making bushels of money, and hopefully the ad payers see value as well.

**Where's the find in search?** What's really lacking in the Internet search category is the ability to find what you seek within the context that you want it, and ideally by using a natural language query string. In Google, if you type, "Looking for a white painted bureau with glass handles," you will get about 60,000 possibilities. The list will include everything from white bureaus at an inn in Vermont, doorknobs made out of glass handles, white glass, white paint and so on.

The issue is that 99.9999 percent of the content delivered is not relevant to helping a person get information to make a decision. This might be OK for certain "informational search" sites. But for those in commerce, not helping a person pinpoint what they need means a lost sale. A commercial site that does not answer a query in a way that helps a person make a decision and lets them find what they seek will not make money.

On commercial sites, "content" and the ability to search it within the context of a query is crucial. Also important is how that content is organized on a site, as is how that content can be searched and used within the context of a person's query.

**Pinpoint answers.** Commercial sites need to let visitors use their natural tongue to make a query. On a large electronics Web site, cables that connect one hardware system to another can be a complex item to find. Cables have different end fittings, lengths and quality based on what they need to do.

With a natural language processing discovery engine, coupled with the ability to rec-

ognize full sentences and in the context that a question is being asked, the cables can be pinpointed to achieve the correct answer. Using this type of system, an electronics firm raised its conversion rate of queries to sales to 11 percent. The standard is less than 1 percent.

This type of discovery technology pinpoints answers that people seek, less so the large aggregators of pages without real context that Google and others offer at a price based on ads served. A smart combination that advertisers on Google need to consider is putting a contextual search and discovery engine on their own Web sites to guide visitors and let them make buying decision quickly. This type of business strategy will yield higher revenue and conversion rates for their ad investment on Google.

If you want to try a commercial site that uses a contextual discovery engine to see the difference, go to [www.maplin.co.uk/Sub\\_Category.aspx?Menu=2&doy=6m4#](http://www.maplin.co.uk/Sub_Category.aspx?Menu=2&doy=6m4#) and click on the Maplin Cable Finder with the photo of the blonde saleswoman. Now type in, "I want a cable to attach my laptop to a keyboard" and see what you get.

By providing valuable textual content, a site not only can increase its credibility but also improve its conversion rate. ■