

Home | Fairfield County Business Journal | Scraping below the surface

Scraping below the surface

Alexander Soule | Oct-01-10, 03:40 PM | 3 Votes

Topics: Technology

Do a Google search of that company's own name and you get 1.4 billion hits. Do the same on Semantifi Inc.'s own search engine, and you get just five.

Give Semantifi CEO Shree Pragada 14 months, and he promises you will find things on Semantifi.com you never would on Google – assuming the Stamford-based startup stays a step ahead of the Silicon Valley giant.

Semantifi's innovative approach to search technology is already drawing eyes, with the company the lone Connecticut startup to participate in last month's DEMOfall 2010 conference in Santa Clara, Calif., among the most closely watched high-tech confabs in the world.

The company's portal is designed to crack open access to what search engine companies dub the "deep Web," which accounts for 99 percent of the information on the Internet but which is largely passed over by Google and Bing's Web crawlers.

Adapting the Wikipedia model
The term "deep Web" is attributed to Michael Bergman, who published a 2001 paper in the Journal of Electronic Publishing that analyzed the problem. Nearly a decade later, Semantifi thinks it has found an answer in the Wikipedia model, in which users contribute their own input.

In the same way anyone can create an encyclopedia article on Wikipedia, anyone can link an accessible database to Semantifi. Hosted by government agencies, businesses, or individuals, many of those datasets escape the attention of traditional search engines, according to Pragada and other deep Web devotees.

In theory, the Wikipedia model should allow Semantifi to extend its reach faster than a small team of Web programmers could do on their own as they attempt to keep up with a constant stream of new sources of data.

The question becomes how effective the company can be. A late August search of the word "Semantifi" on Google dredged up 11,800 returns; when the company's name was entered into its own website, it returned a message reading "Unable to recognize query. Please reenter."

Attracting posters, content quickly is key
Pragada acknowledges the importance of the company adding a large number of datasets in a short period of time, to give early users success using the website even as the company and outsiders add more. He adds that the quantity of search returns served up by Semantifi is less important than the relevance of those returns, which the company is attempting to improve through software innovations.

For instance, searches on Bing and Yahoo using the term "Semantifi" dredged up 8 million results, though cluttered with similar sounding names and words, such as "semantic" and "Symantec Corp." In much the same way, Pragada noted, searches on standard search engines using more-complex phrases often miss the mark, forcing users to scroll through page after page, or to refine their search terms in differing ways – sometimes with success, sometimes not.

Myriad companies have worked on the problem of applying advanced semantic capabilities to database searches, but few have attempted to build a Web search company from that effort. Google itself reportedly has internal teams attempting to improve the company's access to the deep Web.

Founded as ExeCue in 2005, since then Semantifi has raised nearly \$5 million in financing, with its backers including the Connecticut Innovations Inc. venture capital fund backed by the state.

Pragada attributes his company's presence in Stamford to Connecticut Innovations' early support – how fast it will grow is completely dependent on how quickly it can get the broader universe of Internet users linking in data.

Wikipedia, at least, has provided the company with a roadmap to how that can be done.

"Here, we are at least learning from the other revolutions," Pragada said.

Facebook | Twitter | LinkedIn | StumbleUpon | Dribbble | RSS | Email this page to a friend.

0 Comments

Add Comment

Name

Email (will not be published)

Personal Web site/Blog URL

Subject

Scraping below the sur

Comment

Submit

Articles by This Author

Alexander Soule

Exxon seeks OK for \$1B floating gas terminal

Dec-23-07, 07:00 PM

Exxon Mobil Corp. plans to seek approval for a floating liquefied natural gas terminal 20 miles off the coast of New Jersey.

Retail recovery ramps up

Sep-17-10, 09:54 AM

Urstadt Biddle and Macerich say the outlook for retail leasing is improving.

United Rentals shares fall after sale fails

Nov-25-07, 07:00 PM

Shares of United Rentals Inc. fell 25 percent after the company confirmed reports that Cerberus Capital Management L.P. called off its \$4 billion acquisition of the company.

Pink slip issue

Jul-13-08, 07:00 PM

Towers Perrin, a Stamford-based company that ranks among the most prestigious human-resources consultancies in the country, is being sued by two former managers who allege the company violated federal labor laws in firing them following crippling hospitalizations.

Dot-com shipper is dot-gone

Jan-06-08, 07:00 PM

Baker and Taylor inc. of Charlotte, North Carolina has sued Alphacraze.com for \$2.7 million and additional interest and legal costs.

Judge socks UBS with bond ruling

Sep-18-09, 09:44 AM

A state judge ordered UBS to set aside more than \$35 million as a potential award to Stamford-based Pursuit Partners.

Sectra exec dies in crash

Nov-20-09, 09:56 AM

Sectra executive John Goble died in a helicopter crash.

Building materials through the roof

Apr-06-08, 07:00 PM

Construction commodities outpace inflation

PepsiCo battles sales decline, PBG

Jul-19-09, 07:00 PM

PepsiCo Q2 revenue was down 3 percent from a year earlier to \$10.6 billion.

MX Energy no. 6 among 'Hot 100'

May-25-08, 07:00 PM

A new report named MX Energy Inc. the sixth fastest-growing company in the country this decade, after it reported \$704 million in revenue last year from \$200,000 in startup funding in 1999.

Back in Black Top

Mar-08-09, 07:00 PM

Road warriors give new life to old roads

On a chilly day near the Shelton riverfront, the operators of a front-end loader dumped a load of pavement debris into an orange hopper.

Report: 60 Health Net jobs cut

Jul-09-10, 01:38 PM

Health Net of the Northeast Inc. reportedly cut 60 jobs in Shelton.

Energy report

Sep-02-10, 01:58 PM

The Connecticut Clean Energy Fund is offering a new loan program for emerging technologies.

Little credit given job-creation tax incentive

May-06-07, 07:00 PM

As Connecticut's job-creation tax credit nears its one-year anniversary, just one company has given the measure credit for creating jobs in the state.

Workers' comp hikes coming

Nov-20-09, 10:02 AM

Most categories of workers compensation insurance will rise in 2010.

Mountain of money in low-cost loans

Mar-22-09, 07:00 PM

As the U.S. Department of the Treasury attempts to kick-start the economy, officials are counting on loan guarantees and fee waivers administered by the Small Business Administration.

Bank profits slide in 4th quarter

Mar-23-08, 07:00 PM

New data declares that three more commercial banks have reported losses in New York state.

IBM buys data warehouse maker

Oct-01-10, 03:09 PM

IBM is acquiring Netezza, whose systems analyze large volumes of data.

Job losses by default?

Jul-29-07, 07:00 PM

Tighter credit standards may fuel debt defaults, industry observers warn

Even as Fairfield County enjoys a hiring boom with an increase in June of 8,000 new jobs, financial experts are voicing fears that tightening credit standards may spark an increase in business debt delinquencies and defaults in the second half of 2007 – possibly leading to future job losses.

Insurance penalties spike in third quarter

Nov-09-08, 07:00 PM

Connecticut escalated collections due to insurance-industry infractions in the third quarter, on the eve of a new law's enactment that brings exponentially stiffer penalties going forward.



POPULAR | LATEST

1. Three county companies growing fast
2. Small Business Jobs Act offers tax savings
3. Downtown magnetism
4. Seeing green
5. Retailer takes over former news building
6. CRT begins recruiting financiers
7. FDIC puts Wilton Bank on watch list

News by Topic

- Advertising
- Architecture
- Arts
- Aviation
- Banking
- Colleges & Universities
- Construction
- Credit Unions
- Economic Development
- Education
- Green
- Health Care
- Hospitality
- Hospitals
- Insurance
- Law
- Leisure
- Marketing
- Nonprofits
- Real Estate
- Restaurants & Food
- Retail
- Sales
- Schools
- Small Business
- Taxes
- Tourism
- Transportation
- Travel

Reach more decision makers.

09373433

BUSINESS CONTACTS AND GROWING

Search for free

Find Contacts

Or earn credits

Trade Contacts

provided by bizjournals

What is the Best ERP?

Compare Leading Accounting Software Packages

COMPETITION: INTENSE

of the trade

FREE Report

30 Questions Every CFO Must Ask About the Price of Accounting Software

Generations

CLICK HERE - SUBSCRIBE NOW

Westchester Generations

tune in 100.7 WHUD

Quick Quote

Calculate the Cost of Microsoft Dynamics GP ERP Software

of the trade

Facebook logo