

NYS Innovation

This is a Blog to promote the best of NYS Innovation and what is happening in academic research centers and small companies throughout NYS. We will discuss breakthroughs in Info Tech, Nanotech, Semantics, Web 3.0, Biotech, HPC, Advanced Materials, AI, Microelectronics, Drug Development, and all the people who make it all happen.

Home

Thursday, September 16, 2010

Using Human Effort to Make Semantic Search Useful

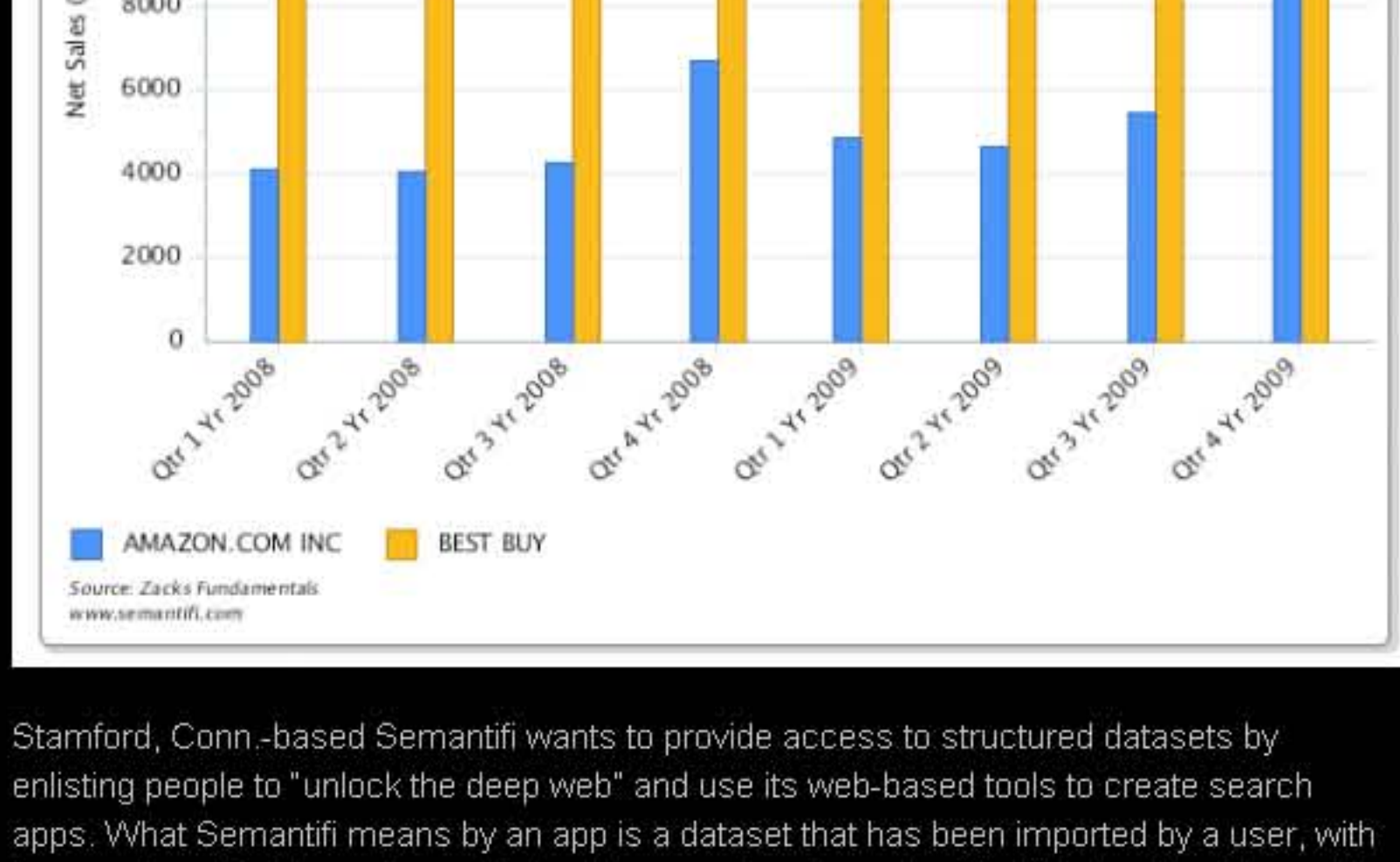
Using Human Effort to Make Semantic Search Useful: *

The best way to find the needle you're looking for in that haystack is to organize the hay. We need to understand how pieces of information relate to each other, or they're not useful. Two new startups I spoke with recently — Semantifi and FindTheBest — are turning big, opaque datasets into online databases that can be queried. Both use human power to format thousands of topic-specific web applications, which help users manipulate data to find more complex and satisfying answers than what could be found with a regular search query.

Semantifi was one of the cooler companies I saw at the DEMO Conference this week. You can use the company's SEC filing search app query "Best Buy Amazon sales last 2 years" and within seconds, you get a table comparing quarterly sales data for each company, as well as an automatically formatted graph (pictured below).

Recent Research

- [Transient Apps: The Consumer Influence on Enterprise Mobility, Part 2](#)
- [Rogue Devices: The Consumer Influence on Enterprise Mobility, Part 1](#)
- [Apple's Path to the Living Room](#)
- [SkypeIt: Skype's Platform for CE-Based Communication](#)



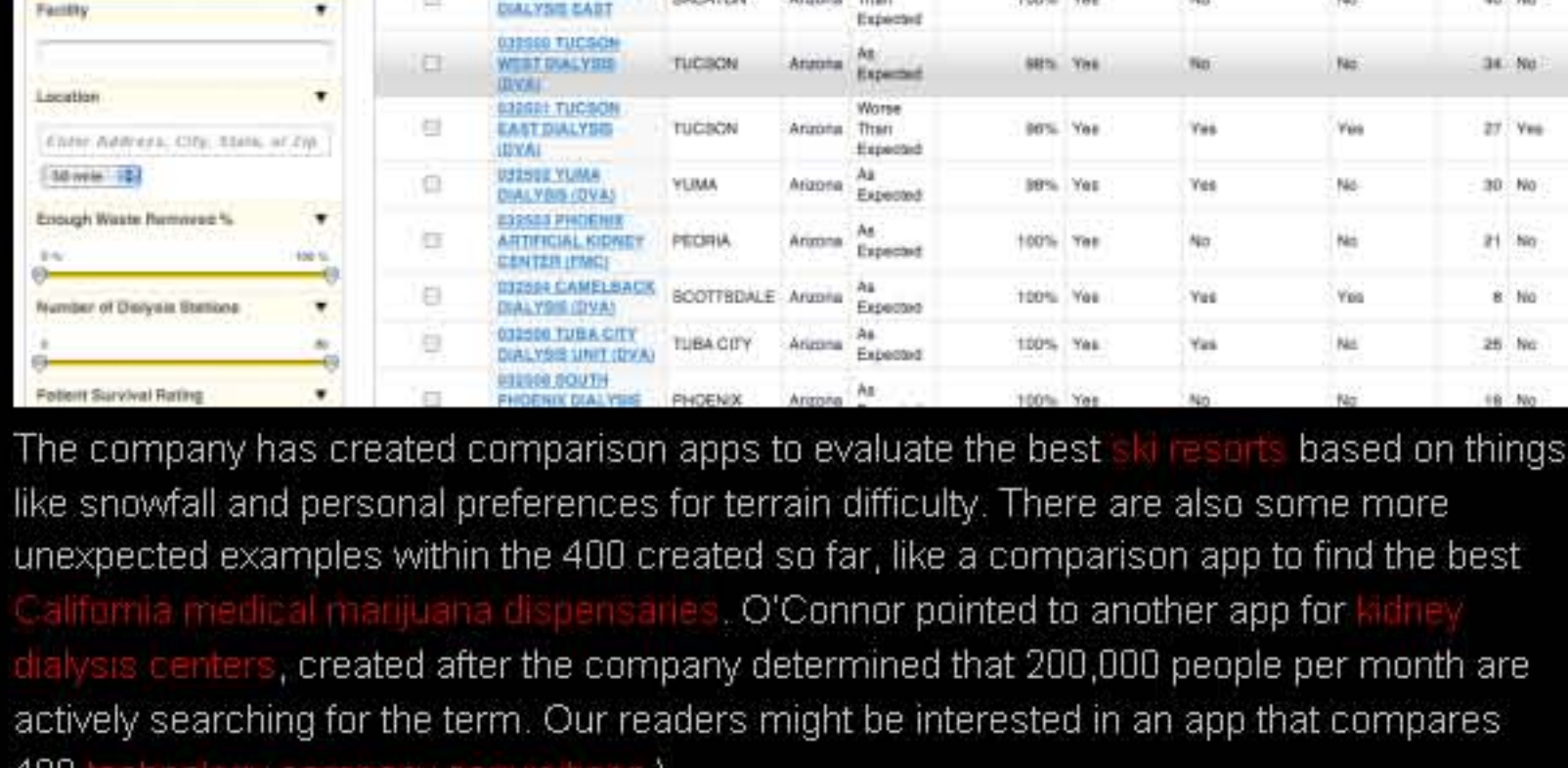
Stamford, Conn.-based Semantifi wants to provide access to structured datasets by enlisting people to "unlock the deep web" and use its web-based tools to create search apps. What Semantifi means by an app is a dataset that has been imported by a user, with the user teaching the system which cells are categories and a little bit about how they relate to each other. Semantifi has seeded its site with search apps to better access government and financial data.

Semantifi, which has raised \$3.5 million in funding, is trying to enlist users to organize data, but it's also striking partnerships with publishers, for example, Zacks Investment Research is helping create apps to query financial data. Publishers can [turn their content and data into a platform](#) and choose to only allow paying subscribers to access their apps.

How might a regular user to participate? CTO Vishy Dasari showed me how to create an app from Amazon product data (taken from a chart provided to affiliate partners) to get answers like "silver digital camera under \$400." If that's the question you're looking to get answered, it would take you many clicks within Amazon's own site search and results pages.

Meanwhile, **FindTheBest's** (FTB) first notable feature may be its founder and CEO, Kevin O'Connor, who also co-founded the ad network DoubleClick, and is now returning to entrepreneurship after a 10-year hiatus. "On the Internet, you can find any piece of information, but you actually want to compare things and make a decision it's very difficult," said O'Connor in a telephone interview. His Santa Barbara, Calif.-based, bootstrapped company wants to help users make comparisons by giving them detailed charts made by its in-house researchers and contract workers.

FTB determines what new topics to tackle by looking at search data. That part sounds a bit like **Demand Media**, but rather than articles or videos, the company creates living databases that look a lot like a Kayak search results page — big chart in the middle of the page, sliders on the left to tweak what the chart is showing, and options to dive deeper and compare a few top items.



The company has created comparison apps to evaluate the best [ski resorts](#) based on things like snowfall and personal preferences for terrain difficulty. There are also some more unexpected examples within the 400 created so far, like a comparison app to find the best [California medical marijuana dispensaries](#). O'Connor pointed to another app for [kidney dialysis centers](#), created after the company determined that 200,000 people per month are actively searching for the term. Our readers might be interested in an app that compares [400 technology company acquisitions](#).

FTB stands out for how manually it's created. If the hope for the so-called "semantic web" is to think more like a human, we should stop trying to conjure magic software and hire some people. The company has 15 full-time employees, eight interns and 20 outsourced workers between its research and product teams, O'Connor said. Any user can edit an app, but the change has to be approved by FTB editors.

Since new search destinations almost never gain market share, it will be important for both these companies to bring users in through searches on Google and elsewhere. They're both actively working on search engine optimization. If you search Bing for "housing starts unemployed last 60 months," a Semantifi page is one of the top results. O'Connor said FTB is focused on attracting searchers for long-tail queries, then exposing them to a broader range of comparison information than they might have known was available.

These aren't the only companies focused on helping people create web databases. I recently wrote about [Needbase](#), a fascinating side project of the travel company ITA that Google just acquired, and Google also just [launched MetaWeb](#) and its open database Freebase. Other competitors include Wolfram Alpha, Socrata and Factual.

These startups answer Tim Berners-Lee's call to turn "raw data" into "linked data." He made a [passionate plea](#) at TED last year to bring data online in order to mash it up "for multidisciplinary purposes, like combining genomics data and protein data to try to cure Alzheimer's." Having what you're doing blessed by the creator of the web is a pretty good start.

Posted by Michael Ridley at [Thursday, September 16, 2010](#)

0 comments:

Post a Comment

Comment as: [Select profile](#)

[Post Comment](#) [Preview](#)

[Newer Post](#)

[Home](#)

[Older Post](#)

Subscribe to: [Post Comments \(Atom\)](#)

Twitter

0 Tweets

Search This Blog

About Me

Michael Ridley
Albany, New York, United States

I am the Director of the High Performance Computing Program for the NYS Foundation for Science, Technology and Innovation. I love technology, politics and all the spaces in-between.

[View my complete profile](#)

DISCLAIMER

The views and opinions expressed herein are those of the author and do not necessarily reflect the opinion of the author's employer.

Subscribe To NYS Innovation

Posts

Comments

Sociable

- [Blog](#)
- [Facebook](#)
- [Google Bookmarks](#)
- [Google Plus](#)
- [LinkedIn](#)
- [Slide](#)
- [Twitter](#)
- [Facebook](#)
- [Flickr](#)
- [Twitter](#)

Blog Archive

- ▼ 2010 (1691)
 - 10/03 - 10/10 (80)
 - 09/26 - 10/03 (173)
 - 09/19 - 09/26 (158)
 - ▼ 09/12 - 09/19 (128)
 - [Wall Street Journal: Netfix, Google to pressure Cable](#)
 - [Robotix Technology Morphs into Assistive Exoskeleton](#)
 - [When will the first Earth-like planet be discovered?](#)
 - [Channeling crowdsourcing into astronomical work: ftt](#)
 - [Google testing higher-quality video chat in Gmail](#)
 - [Research Team Assesses Environmental Impact of Org. 2010 is the year of participatory platforms: ftt](#)
 - [Rochester Fabricator opens in W Seneca: ftt/pjn](#)
 - [IBM adding to software portfolio: ftt/ping:fm/59](#)
 - [White House labels fiscal stimulus projects 'sequest'](#)
 - [Cybercrime is world's most dangerous criminal threat](#)
 - [Doubling Lithium-ion Battery Storage: ftt/ping:fm](#)
 - [Credit Card with a Computer Inside: ftt/ping:fm](#)
 - [Best Buy CEO: iPad Cannibalizing U.S. To 50 Percent](#)
 - [Lawlist targets advertiser over sneaky HTML5 ppc/ad](#)
 - [Why isn't There A Central Database Of E-Voting Pkts?](#)
 - [Radiation machine could cut cancer treatments in 1/2](#)
 - [I.B.M.: Africa is the Next Growth Frontier: ftt/p:fm](#)
 - [Verizon LTE in 30 cities by year end, AT&T aims to](#)
 - [The 5 best and 5 worst private equity funds: ftt/p](#)
 - [Extreme Netfix: Some Users Have Watched 50,000 Movies](#)
 - [GENUS Opped to draw young researchers to Oswego](#)
 - [New Types Of Rock Found On Moon By Researchers At](#)
 - [Kodak Celebrates 125 Years of Innovation at Photo](#)
 - [Comcast/Connect social network goes live: ftt/p:fm](#)
 - [Buffalo private-sector job count rises: ftt/p:fm](#)
 - [Alcoa Joins Change the Equation to Solve America's](#)
 - [Open source Facebook replacement Diaspora drops in](#)
 - [AT&T to launch LTE by mid-2011 \(PewGoldstein:fm\)](#)
 - [Verizon Wireless to Launch 4G LTE Service in 30 U.S.](#)
 - [Google gets into AI with an API - iProgrammer: ftt/p](#)
 - [Social Networking Plays Bigger Role in Job Searching](#)
 - [Using Human Effort to Make Semantifi Search Useful](#)
 - [LTE to Boost Demand for Mobile Bandwidth, Network](#)
 - [IBM again No. 2 among top 100 brands, Cisco also in](#)
 - [Startup aims to stream Flash games to iPads: ftt/p](#)
 - [RPCI nurses honored: ftt/ping:fm/ACZcS](#)
 - [Presidential Advisory Group Makes STEM Education R](#)
 - [What's faster than rural internet upgrades? Carrier](#)
 - [No one trusts the SEC: ftt/ping:fm/ANJPC](#)
 - [Time Warner Cable Really Hates Online Video: ftt/p](#)
 - [Yale Researcher Develops "Supercomputer" That Mimic](#)
 - [Halo: Reach tallies \\$200 million on launch day: ftt](#)
 - [True Voice Technologies introduces Interactive Spe](#)
 - [The Two-Stroke Engine Reconsidered: ftt/ping:fm/nyupdy](#)
 - [Grooming NYC's Top Executives to Go Gre](#)
 - [In the heart of a wind turbine: ftt/ping:fm/Vtdm](#)
 - [IBM Buys OpenPages To Spice Up its Business Analyt](#)
 - [US Neurosurgery Department Ranks 7th in North Amer](#)
 - [The \\$200M challenge to achieve energy independence](#)
 - [Move Gets Streaming Patent, Are Adobe & Apple Hoss](#)
 - [New database could speed up drug discovery: ftt/p](#)
 - [AIG Retains Its Risk Approach: ftt/ping:fm/AbdH4](#)
 - [iOS 4.2 turns iPad's screen lock into volume switc](#)
 - [Smart grid cyber security: initial set of guidelin](#)
 - [IT operations in a cloudy word: ftt/ping:fm/az55](#)
 - [Google Music: Part iTunes, Part Lara, Part Awesome](#)
 - [Blog - Friends Provide an Epidemic Early Warning in](#)
 - [Senators Make One Last Push For Bad Patent Reform](#)
 - [Bing could get access to anonymized Facebook data](#)
 - [T-Mobile to implement a Twitter Tax on Texts? \(CJL\)](#)
 - [The Web 2.0 Summit Program Is Live: ftt/ping:fm](#)
 - [Scientists reveal battery behavior at the nanoscale](#)
 - [Blog - Intel Puts Physics in the Cloud to Give Fun](#)
 - [Kodak Showcases Award-Winning Packaging Solutions](#)
 - [Lighttower expands in N.Y. with buy: ftt/ping:fm/v](#)
 - [U.S. tech industry adds jobs, remains behind last](#)
 - [Microsoft Launches Massive Wireless Hotspot \(Estie\)](#)
 - [Microsoft Launches Massive Wireless Hotspot \(Estie\)](#)
 - [Nielsen: Bing Passes Yahoo To Become #2 Search Eng](#)
 - [Broadband duopoly leaves ISPs competing on speed](#)
 - [Broadband duopoly leaves ISPs competing on speed](#)
 - [Linux Torvalds becomes an American citizen \(Paul M](#)
 - [NVIDIA Releases Parallel Insight 1.5, CUDA Toolkit](#)
 - [IBM Chief Slams H-P: ftt/ping:fm/wOeag](#)
 - [Euro Disney to Expand: ftt/ping:fm/FTC17](#)
 - [Coining Warnings of Weakening Glass Ceiling: ftt/p:fm](#)
 - [Verizon Finds an Opening in Japan's IT Market: ftt/p](#)
 - [Few Online News Use Growing But Traditional Metho](#)
 - [How your cell phone camera can improve the perform](#)
 - [SyracuseCoE Symposium 2010 presents 'Restoring Sus](#)
 - [The iPad: A100 - A Worthy iPad Competitor? ftt/p](#)
 - [Why Does Everyone Underestimate Netfix? ftt/p:fm](#)
 - [iPod really starts dominating on iPads: ftt/ping:fm](#)
 - [Google TV officially launching October 17th? \(Ben](#)
 - [Google Research Carves Space-Time to Optimize Vide](#)
 - [eBay: Our merchants are good for the earth: ftt/p](#)
 - [Sprint races to launch 4G in D.C. ahead of Verizon](#)
 - [Wheelchair Makes the Most of Brain Control - MIT T](#)
 - [Medical Messenger Uses HODAR Captura Pro Software](#)
 - [Sensor Networks Top Social Networks for Big Data in](#)
 - [Consciousness and Computer Vision: ftt/ping:fm/ft](#)
 - [New Top 500 list: instrument gives Albany, NY a nationa](#)
 - [SUNY Canton's College in Bosnia Recognized by Pres](#)
 - [Grant to fund SUNY Oswego study of women in scienc](#)
 - [Sciences building projects to include revamped Ric](#)
 - [NSF supports safer internet project: ftt/ping:fm](#)
 - [Researchers attack plum pox to save N.Y. fruits: ftt](#)
 - [Google Says Google Reader Is Doing Just Fine](#)
 - [Disney Hikes On HTML5 \(Dianne de Guzman/Silicon AI](#)
 - [Video Games Lead to Faster Decisions that are No L](#)
 - [President Shirley Ann Jackson and Four Professors](#)
 - [Albany Helps NYS on High Speed Internet Project in](#)
 - [Stena creates business incubator: ftt/ping:fm/6h](#)
 - [ADA Technologies gets grant for prosthetics interf](#)
 - [Disney World As Entertainment: ftt/ping:fm/20Fqg](#)
 - [Economy So Bad, People Aren't Even Buying Weapons](#)
 - [Wheelchair Makes the Most of Brain Control - MIT T](#)
 - [Mountain View's Computer History Museum builds \\$18](#)
 - [Mountain View's Computer History Museum builds \\$18](#)
 - [Sprint races to launch 4G in D.C. ahead of Verizon](#)
 - [Autodesk gives startups cleantech tool: ftt/ping:fm](#)
 - [Autodesk gives startups cleantech tool](#)
 - [State of the Schools: Education key to future of f](#)
 - [State of the Schools: Education key to future of f](#)
 - [Economy So Bad, People Aren't Even Buying Weapons](#)
 - [YouTube reportedly testing live video streaming in](#)
 - [Trent signs it's truly a jobless recovery](#)
 - [Get on the Optical Bus - IBM's light-powered links](#)
 - [State awards \\$108 million to expand use of health](#)
 - [Comcast's NBC-U.Dreams May Be Online Video's Night](#)
 - [Revenge of the nerds: Should we listen to futurist](#)
 - [Apple may abandon Intel's WinMn \(DAB in need\)](#)
 - [Patent Office Admits the Truth: Things Are a Disas](#)
 - [How Google Reader and RSS beat Twitter in some cas](#)
 - [Reading E-Books in All the Colors of the Rainbow \(](#)
 - [Leaked: Acer Laptop with Dual Touchscreens \(Alex G](#)
 - [FCC move to release White Spaces has tech firms dr](#)

Picture Window template by Josh Peterson. Template images by enot-poloskun. Powered by Blogger.