Yahoo Aims To Be Research Powerhouse

By Wade Roush October 12, 2005
When U.S. Web surfers are searching for information, 80 percent of them turn to one of three sites: Google, Yahoo, or Microsoft's MSN. And because the keyword-specific ads displayed alongside search results command such a premium, these three leading search companies have spent the last two years sparring relentlessly with each other, adding free services -- desktop toolbars, blogs, gigabytes of e-mail storage -- designed to turn casual users into loyal ones.

But when it comes to developing new Web technology, one of these companies has lagged behind the others. Surprisingly, it's not Microsoft, but Yahoo.

At Google, research is woven into the fabric of the company: software engineers are required to spend 20 percent of their time on far-out ideas, a policy that's given rise to a host of spin-off Google sites.

Microsoft, for its part, has funded extensive research in areas such as data mining and information retrieval, including a system that assembles information from the Web and a user's hard drive before he or she has even realized they need it.

But for Yahoo, having a research operation that helps to invent emerging information tools has never been a major priority. Indeed, until two years ago, the company didn't even have its own search engine -- it rented Google's.

But now that's changing -- and fast. In July, Yahoo hired Prabhakar Raghavan, the former chief technology officer at enterprise-search provider Verity, to lead its 40-person research division in the company's Sunnyvale, CA headquarters.

Raghavan, who is also editor-in-chief of the *Journal of the Association for Computing Machinery*, has proceeded to put Yahoo Research on the map by wooing top researchers, such as Andrew Tomkins, a text-analytics expert so well-regarded for his work on Web buzz-tracking at IBM's Almaden Research Center that Fortune magazine called him one of IBM's "golden geeks." More hiring announcements are imminent, too, according to Usama Fayyad, Yahoo's senior vice president and chief data officer.

"What Prabhakar has achieved in two and a half months of recruiting is to attract the top minds in the world in the spaces of search and social media," says Fayyad.

Raghavan was more modest in a three-way conversation with *Technology Review* last week. But Fayyad said he was quite serious about rapidly making Yahoo a major player in new Web-based tools.

Some outside the company caution that building a strong research division takes time, though. "Research is a long-term investment in the future, not a way to fix the present," says Rick Rashid, Microsoft's senior vice president of research and the former director of Microsoft Research. "Based on what I have seen in the past, I think it is fair to assume that it will take a number of years for the impact of a new research organization to be felt in earnest by a company.

"At the same time, there is great value in hiring good researchers and bringing them into a corporate environment," Rashid says, "and that value can translate quickly into an influx of ideas and perspectives that can change the nature and quality of a company's products even in the short term."

The turnaround at Yahoo came, according to Fayyad, when Yahoo's leaders began to ask whether the company's existing research effort would help Yahoo take advantage of the changing nature of the Internet, particularly the growth of interactive Web-based applications such as blogs, auctions, and travel planning.

"We had an early version of Yahoo Research where we were very pleased with the performance, but we weren't very happy with the vision of where we were going," Fayyad says. "A huge proportion of the advertising dollars in the world are going to flow [toward interactive applications]. We need to know, 'What are the real underlying problems?'"

Raghavan says the work of the research division is being organized into five focus areas:

**Search and retrieval**, including studies in what Raghavan calls "adversarial" information retrieval. "There is this ongoing, two-sided game between the Web search engines on one side and the spammers on the other," he explains. "You have to think two moves ahead of the spammer. You have to ask, 'If I make these fixes, how is how the fixes would manifest themselves, and the spammers are going to be able to infer this and that, so I better do this too.' It's an evolving game that's never going to end."

**Data mining and machine learning.** One illustrative project is Mindset, which is available for testing at next.yahoo.com, the company's public online area of research projects in progress. Mindset puts sliders alongside Yahoo search results that allow users to tailor search
rankings to highlight either editorial or commercial content. Machine-learning algorithms use keywords and other clues inside Web pages to gauge their score on this axis. In principle, Mindset-like sliders could be used to adjust search results along any parameter.

*User interfaces and user experiences*, including social experiences using short-range wireless protocols such as Bluetooth for locating friends and potential contacts. The question that interests Yahoo in this area, according to Raghavan, is "How do you have a million or a billion people interacting with a million or a billion computers and other devices and create something better than a million individual experiences."

*Utility computing*. Raghavan's definition of the problem: "How do you spread a million PCs around the world and have them work as one big mainframe, and extend that beyond the PC into the [cell-phone] handset."

*Microeconomics*. Economic theory can illuminate many types of user behaviors and beliefs on the Web, says Raghavan. He says Yahoo would like to understand how to deploy incentives to reward honest users of social media such as *My Web 2.0* and *Flickr*, sites where users can share Web sites and photos they've "tagged" with their own keywords. In October, *Technology Review* named Yahoo researcher David Pennock as one of the 35 most promising innovators under age 35 for his work on predictive markets such as Yahoo's *Tech Buzz Game*.

Although the research in all five areas is ultimately aimed at making Yahoo a more popular destination for Web users, not every project at Yahoo Research has a direct relationship to an existing product.

"There isn't in all cases a one-to-one mapping between our research areas and the business units that Yahoo has," says Raghavan. "But that's as it should be. We see ourselves not just as tactical problem solvers for this or that business but really as architects for the businesses in the making."

Microsoft's Rashid says he has two suggestions for Raghavan and Fayyad: "The most critical things you can do when you start a new organization are, one, hiring the right people, and, two, establishing the right values. If you don't exemplify the right values even the best people will fail. If you don't have the right people, no amount of organizational support will help."

Establishing the right values often takes time. Meanwhile, Yahoo's Raghavan is searching for the most valuable researchers.

"Developers involved in a software organization typically produce $1 million in revenue per year, per developer," Raghavan says. "If my scientists each produce $1 million in revenue, that's not very interesting. The thing I look for is for our scientists to think about hundred-million-dollar ideas or billion-dollar ideas."

50.1061446927654

CATEGORY_KEY WORDS: Telecom/Internet