

washingtonpost.com

New Services Are Minding Your Business

By Leslie Walker

Thursday, March 4, 2004; Page E01

Admiral John Poindexter must be irked. The furor over his "Total Information Awareness" surveillance plan helped trigger his departure from the Pentagon's research agency last year. Now entrepreneurs are selling new data-mining tools that do similar things on a smaller scale -- and venture capitalists are hurling money at them.

The new tools are business cousins of Friendster and the other social-networking sites that took the Net by storm last year, helping people set up electronic networks of friends. But like Poindexter's ill-fated scheme -- a massive plan that would have allowed the creation of electronic dossiers on millions of Americans to sniff out potential terrorists -- the business-networking tools raise prickly privacy issues.

Rather than creating simple electronic Rolodexes, business-networking tools go deeper. They aim to help companies make faster sales by identifying and mining electronic "new-boy networks" of personal contacts. They help professionals get introductions to people they don't know, relying on mathematical formulas to analyze data crumbs left by people's electronic communications.

On the surface, products from start-ups such as Visible Path Corp. and Spoke Software Inc. may seem helpful and innocuous, but in the background many users may have no idea how much data is being mined.

I was stunned when I signed onto Spoke's network for the first time this week and searched for my name. Up came an eight-year-old photograph of me, accompanied by a 150-word profile that I had hastily written years ago. It sits on an old Web page that is no longer publicly linked on The Washington Post's Web site. My "dossier" also linked to five seemingly random articles I wrote in years past. When I searched on other people's names, it showed me their dossiers, too, along with a list of people who presumably had in their contact lists both my e-mail address and those of my search targets. What I found interesting -- and chilling -- was how automated this obviously was, considering I gave Spoke no information except my e-mail address when I signed up.

Vendors seem acutely aware of the potential to invade privacy, especially since some of their tools do massive snooping on both internal corporate networks and the public Web to generate "dossiers" of people who never signed up to participate.

"We have hired a privacy officer to address these concerns," said Ben Smith, chief executive of Spoke, based in Palo Alto, Calif. "And if you want to be off our system, we have a process that allows you to opt out. So far we have removed two out of the 15 million people on our network."

You read that right: 15 million people have been identified so far by Spoke's automated profiler. That's a lot, considering Spoke went live in October 2002.

Both Spoke and Visible Path sell their software primarily to corporations, not individuals. The idea is to provide tools for finding potentially helpful business partners and making blind introductions -- allowing, say, a lawyer for Silly Computers Inc. to electronically ask a former classmate from Harvard who once did legal work for Microsoft to help him pitch a business deal to Bill Gates.

The potential introduction would be "blind" in the sense that the Silly Computer lawyer wouldn't know the identity of his former classmate. After comparing a gazillion electronic address books, the software would identify someone who had swapped e-mail with Gates and also had a career overlap with the lawyer. It would offer to forward the lawyer's message to his ex-classmate without telling the lawyer the intermediary's name.

If the potential upside is that the tools help workers squeeze more out of everyone they've met and make new contacts, the potential downside is, well, the same. A world in which salespeople are using higher-powered electronic networking tools would inevitably be a world in which more pushy people engage in more intrusive behavior.

Both Spoke and Visible Path say they provide controls to let users decide how much personal information they want to be visible across the network.

"Our system assures that neither their relationships nor their contact information is ever shared without their explicit approval," said Antony Brydon, chief executive of Visible Path.

The founders of Visible Path and Spoke are convinced corporations will buy their software once they see how it can boost sales. Spoke reports having 30,000 users in a dozen large companies and two dozen smaller ones. Smith said the tools appear to be having impact -- in some cases, doubling the number of leads that result in actual sales.

How does it work? Both Spoke and Visible Path send "crawlers" around a corporation's internal computer network -- sniffing telltale clues, say, from employee Outlook files about who they e-mail and how often, who replies to particular messages and who doesn't, which names show up in electronic calendars and phone logs. Then it cross-references those snippets with information from other company databases, including sales records from PeopleSoft and SalesForce.com.

Spoke goes further, comparing data from internal corporate networks with stuff on the public Web. It applies algorithms to map and rank relationships employees seem to have with co-workers and outsiders. Spoke works a bit like Google's crawler -- only instead of Google's "page rank" formula for determining the relevance of a page to keywords, Spoke has a "closeness index" that measures the strength of people's connections.

Visible Path doesn't search the public Web, but it also attempts to measure how well people know one another. It scours any database a corporation is willing to make available, starting with e-mail, address books and instant messaging systems. With e-mail, it ignores the content of message and instead looks at such details as how quickly people respond to particular senders; the ratio of messages they send back and forth (who sends more?); how many blind carbon copies are sent; and the duration and momentum of communications.

"There is actually a lot of science on this stuff," said Smith, noting that social scientists have explored networking theories for years. "The problem with the science before was there wasn't enough data. Now we are testing this theory with very large, very available data sets that did not exist before."

Smith said he wasn't a believer when someone first proposed the idea to him. "I told them I thought it was goofy."

That was before Sept. 11, 2001. Afterward, Smith found himself doing work for national security agencies that involved data trails left by terrorists. One glimpse of a massive map of electronic connections between suspected terrorists changed his mind. "I saw you could make amazing things happen," he recalled.

No doubt that's true, but let's hope the truly amazing part doesn't turn out to be some data-Chernobyl, triggered by even more aggressive relationship-mining tools that wind up humiliating rather than helping people.

And by the way, does anyone believe Poindexter's plan is really dead?

Leslie Walker's e-mail address is walkerl@washpost.com.