

## Google on Steroids

### Software company making name for itself in Columbia

By C. GRANT JACKSON  
[Gjackson@thestate.com](mailto:Gjackson@thestate.com)

Software developed by a Dutch company is transforming the way scientific and academic research is done — and the company has put its North American headquarters in Columbia.

Collexis Inc., which is led by Irmo native Bill Kirkland, moved into the Capital Center building in downtown Columbia last week.

The company also announced a collaboration to develop one of its products — the Expert Profiler System — with Johns Hopkins University in Baltimore.

Collexis software searches and mines large amounts of data and information. The company's clients already include such research heavyweights as the National Institutes of Health, Merck, Lockheed Martin and the World Health Organization.

The company is attracting a lot of attention.

"The reality is that we are in discussions with every major information technology company globally about partnering right now," said Kirkland, 44, the chief executive officer.

What sets the software apart from standard search engines is its ability to search and discover concepts — not just simple text.

- In a Google search, you type a few words or a phrase into a window, execute the search and get back thousands of results — many irrelevant to the information you need.
- In a Collexis search, a researcher can paste an entire paper or even a book into the window, execute the search and get an ordered list of only relevant material. Collexis knows the difference between Jaguar the car and jaguar the animal.



TIM DOMINICK/TDOMINICK@THESTATE.COM

Collexis Inc., which developed a high-powered search engine, bases its U.S. headquarters in Columbia and its global headquarters in the Netherlands. From left are Steve Nash, director of professional services, Stephen Leicht, director of sales and operations, and CEO Bill Kirkland.

**Collexis**

[www.Collexis.com](http://www.Collexis.com)

## Fingerprinting Helps Mine Through Loads Of Data At Once

C. GRANT JACKSON, *Business Editor*

Call it Google on steroids - and even that is an inadequate description.

Software developed by Collexis enables instantaneous searching of large amounts of data and information using a process called fingerprinting.

A Collexis search makes use of a specialized vocabulary within a given field, such as medicine or finance.

It uses that vocabulary or thesaurus to then identify concepts that might have different names or forms, but which represent the

The Collexis technology can be employed by any industry or institution that has a specific language, or Thesauri, and that needs to search public databases or its own private databases. The technology is also language-independent.

Collexis is focusing on five key markets:

- University, medical research
- Government
- Health care
- Financial services
- Legal

Each represents a market in excess of \$1 billion, according to the company, with university research, health care and financial services expected to exceed \$10 billion by 2010.

The company is looking at an 18-month road map for development, and taking the company public is a possibility, Kirkland said.

### **ROOTS IN SHARED DEVELOPMENT**

Collexis B.V., the Dutch parent company, was formed in the Netherlands in 1999. The Collexis software has its origins in the European scientific community. Development was funded by the Netherlands Organization for Scientific Research and the German Agency for Technical Cooperation.

The original software was developed to help information sharing among scientists through the SHARED project, a research group dedicated to improving research for Third World development.

Collexis Inc., a wholly owned subsidiary, is principally a marketing and sales operation. For now, all research and development will stay in the Netherlands. Collexis developers work in collaboration with Erasmus University Rotterdam.

"We share the expense of these people, but the intellectual property is owned by Collexis," Kirkland said.

The Columbia operation opened in February in space provided by local investor Don Tomlin, who has helped incubate several knowledge-based companies.

same idea - such as different ways to express a human gene or to refer to a drug.

The content of a piece of text is represented by a fingerprint, a small and unique representation of the text.

Fingerprinting is like highlighting the important passages in a piece of text - and the concepts on either side of the word.

Using fingerprinting, a Collexis search discovers knowledge.

"We refer to it a lot of times as a search engine, but it goes into so much more - almost approaching artificial intelligence," said Bill Kirkland, CEO of Collexis Inc., in Columbia.

Collexis allows researchers to see connections between seemingly unrelated items. The search then orders or indexes the results by relevance and importance.

"That natural process that takes place over a two- or three-year period, the system automates and you can see all the concepts, what the overlapping relationships are, what the strength of the relationships are," said Stephen Leicht, director of sales and operations.

"It just kind of thinks like a researcher does and it does it very quickly," Leicht said.

Collexis' speed is one of its amazing attributes. Search results are displayed instantaneously.

"We haven't talked to anybody who hasn't said, 'Wow!'" said Steve Nash, director of professional services for Collexis.

Collexis' first product for the North American

## **FINDING A HOME**

Collexis is in Columbia because Kirkland wanted to come home.

When the chance to develop the U.S. arm of Collexis came along, Kirkland was being courted by Atlanta and took calls from Florida.

"I said if we are going to do it, let's do it in Columbia," Kirkland said.

The former U.S. head of IBM's health care and life sciences Business Consulting operation is an Irmo native and USC graduate. He is the son of the late Henry Kirkland, a well-known attorney.

Collexis Inc. initially was capitalized by private investors, including Mark Germain, a New York investor who has been involved with more than 20 companies in the biotech field.

The Collexis software caught Germain's attention and he called Kirkland, whom he had met while Kirkland was at IBM, and persuaded him to help launch the U.S. operation.

Kirkland put together a small team in Columbia. His key lieutenants are:

- Stephen A. Leicht, 30, director of sales and operations, founded, ran and sold International Telecommunications Distributors, serving as chairman and CEO until late 2001. After the sale, he completed his MBA at Penn State and spent five years at IBM before joining Collexis early this year.
- Steve Nash, 46, director of professional services, is a former Policy Management Systems executive. Most recently he managed internal development and implementation teams for BlueCross BlueShield South Carolina.

Collexis has a team of about 50 people nationwide, 15 of them based in Columbia, Kirkland said. "We are a mobile organization just like IBM."

Sales are handled largely by contractors on commission.

## **FITTING IN WITH COLUMBIA'S DREAM**

market has been "Knowledge Disease Dashboards" for academic and scientific researchers.

A Knowledge Disease Dashboard allows a researcher to search PubMed, the National Library of Medicine's index of biomedical research from around the world, correlating data and identifying trends and discovering relationships among diseases, chemicals, medications and genes.

"It lets you move a whole lot quicker," said Tom Hodge, a University of Georgia senior research scientist. "It sort of thinks like I think when I'm trying to look up stuff to go after."

"They made a (dashboard) for me for HIV and flu, which just hones in on it very specifically, and that thing is just unbelievable."

"It does everything but baby-sit the children at this point."

Beyond the disease dashboards, Collexis is working on several others applications:

\* "Eye for Patents" is the first full text conceptual search for patents. Currently in 11 pilot sites, the software can look across seven global patent databases. USC's Office of Technology Transfer is among the pilot sites.

"We are pulling the ideas out of the text, not just the words," Kirkland said. "How important is that when you are doing a patent search? It is your ideas you want to patent, not the words."

The company also has fingerprinted more than 450,000 global patents in the life sciences.

While Collexis' presence is small, it is the company's growing reputation that is an asset to Columbia as it tries to build a knowledge-based economy.

"In many ways Collexis is the embodiment of what we're doing in Columbia," Mayor Bob Coble said, "because the company's technology is all about making connections — connections between researchers thousands of miles apart, connections among entrepreneurs and academic institutions, connections among all those whose work can make the quantum leaps in knowledge that spur revolutionary ideas and scientific discoveries."

Leicht, the company's acknowledged whiz kid, said companies like Collexis contribute to that dynamic by bringing in people from out of state.

"If you are really going to drive a knowledge economy," Kirkland added, "we and companies like us are how you do it."

The company looked at space in the USC Columbia Technology Incubator, Kirkland said, but "we're not a traditional start-up."

A meeting with Greater Columbia Chamber of Commerce president Ike McLeese led Kirkland to investor Tomlin, who was so impressed with the technology that he immediately offered them space in the Forest Executive Plaza, the building that Tomlin owns at 4500 Fort Jackson Blvd.

"They chose to come here and be here and live here and work here because they like Columbia," Tomlin said. "They could have done it anywhere in the country and it would have been more convenient."

### **HIGH PRAISE**

Thomas Hodge, a senior research scientist at the University of Georgia studying infectious diseases, was one of the software's early U.S. users.

"It lets you move a whole lot quicker," said Hodge, who has no financial interest in the company. "It sort of thinks like I

\* The Expert Profiler System develops a profile of a researcher based on what that researcher has published.

Nobody can do anything like it, Leicht said. Expert Profiler does more than just associate names with documents, he said - it gives you an idea of what the person knows.

Collexis applications can be combined to find particular kinds of knowledge.

"We have profiled over 2 million people just from PubMed," Leicht said. "That is 2 million experts in the biomedical space."

The Collexis technology is also very versatile. All it needs is an existent thesaurus or the ability to build one and access to databases, either public or private.

"It is a horizontal view of the world," Kirkland said.

-Grant Jackson

think when I'm trying to look up stuff to go after. If I've got an idea, I need to be able to look at it logically. That is the way I think. And that is what this thing does for you.

"It is pretty simple, and the more information you give this program, the better it is. It really hones it down. It is one of the best things I have seen in a long time.

David Blake, a former vice president for research at Emory University and vice dean of research at Johns Hopkins School of Medicine, was so impressed with the software that he accepted a position as the chairman of Collexis Scientific Advisory Board.

The software enables a scientist to complete the hours-long process of searching literature in seconds, Blake said.

"Some of the things that Collexis can do, you couldn't do in an infinite amount of time without Collexis," he said.

"You will be so much advantaged in the biomedical sciences with this software that — once people have it — other people are going to consider it a competitive disadvantage not to have it."

By C. GRANT JACKSON  
[Gjackson@thestate.com](mailto:Gjackson@thestate.com)