

Analyze this...

by Joshua Jaffe in San Francisco Posted 10:40 EST, 2, Dec 2005

Four years ago, home appliances manufacturing giant Whirlpool Corp. had to determine quickly how many of its customers had purchased a particular microwave oven so that the company could alert them about a problem with the model. Over a harried weekend, five employees pored over customer call center data, concluding that 14,500 customers had to be reached. After contacting all of them, Whirlpool determined that only 800 people had experienced problems with their microwaves. The blizzard of phone calls was nonetheless considered a customer relations success.

Three years later, Tom Welke, then Whirlpool's vice president for North American customer care, decided out of curiosity to test a new text analytics software program that held out the promise of swiftly crunching reams of unstructured data for key clues. He poured that same raw microwave-oven customer-call-center data into startup software designer Attensity Corp.'s new Discover text analytics program to see how many customers the software would have advised him to contact.

Five minutes later, the program showed that 1,200 customers might have been affected by the microwave malfunction. Welke was so impressed at the speed with which the automated software program came up with the much more accurate figure that he's become one of corporate America's biggest fans of text analytics. "I saw the power of diagnosing unstructured text," he recalls. "I realized I should spend time on this and make this work."

Which he promptly did. Many of the company's contacts with customers are now recorded and analyzed by Attensity's software, which picks apart the text with its patented linguistic technology to identify trends that Whirlpool might not have ever recognized before. That decision paid off earlier this year when the company spied unexpected problems with a new line of dishwashers it had recently unveiled.

After the text analytics software noticed a spike in customer calls one week, the company discovered that 150 people from its hometown of St. Joseph, Mich., were complaining that their Whirlpool dishwashers were suddenly leaving spots on their glasses. It turned out that the local water department had changed the chemicals they used, which affected the hardness of the community's water.

The new text analytics software program, which constantly analyzes customer call records, e-mail feedback and service technician reports, allowed the company to identify the complaint and act quickly to offer remedies. Welke says that without the software those 150 complaints would likely have vanished in the sea of 400,000 customer contacts the company receives each month.

The new software, says Welke, will help ensure every one of those 400,000 complaints can be analyzed to position Whirlpool to compete more effectively and efficiently. Such cutting-edge text analytics, long viewed as too expensive or just not possible, is now emerging as a viable tool capable of helping large organizations make sense of their unstructured data consisting of numbers, letters, words or paragraphs contained in unruly places such as e-mails, phone records, service reports or the Web. Business intelligence vendors such as Business Objects SA, Hyperion Solutions Corp. and Cognos Inc. make software that helps companies prepare, manage and analyze information residing in orderly databases. Yet only about 15% of most companies' data sits in structured database. The rest is unstructured, which is why a crop of text analytics startups such as Attensity and ClearForest Corp. have emerged in recent years. Even large companies such as IBM Corp. and SAS Institute Inc. are trying to gain a piece of this growing market. They all hope to help large organizations use the data to improve customer management, marketing and product development and quality.

Despite the promise, text analytics adoption among corporations and other large enterprises is still very low. Misperceptions about the complexity of integrating text analytics software, a lack of understanding about its capabilities or potential return on investment and the reticence of early adopters to discuss what they view as a competitive advantage are the reasons for the slow uptake, according to research firm IDC Corp., which completed a report on text analytics in June.

Slowly, though, the value of text analytics is becoming more and more apparent in the marketplace. The

software vendors have improved their products so that they are more tightly integrated into existing business intelligence suites, allowing search results to be viewed immediately and in reports that are more palatable than ever before. And the software can be customized so that shorthand specific to a company can be analyzed.

Unstructured bonanza

Companies large and small are developing enterprise applications to analyze and organize unstructured text

Company	Location
Text analytics	
Apollo Data Technologies LLC	Chicago
Attensity Corp.	Palo Alto, Calif.
Basis Technology Corp.	Cambridge, Mass.
ClearForest Corp.	Waltham, Mass.
Nstein Technologies Inc.	Montreal
Data and text analytics analysis	
Inxight Software Inc.	Sunnyvale, Calif.
Business intelligence, data and text mining	
IBM Corp.	Armonk, N.Y.
Megaputer Intelligence Inc.	Bloomington, Ind.
SAS Institute Inc.	Cary, N.C.
SPSS Inc.	Chicago

Source: The Deal

Equally important, research is showing corporations that the return on investment from text analytics can be impressive. Most companies that use the new software are reluctant to discuss the financial benefits in any detail, but technology investment research firm Nucleus Research Inc. in March polled 16 users of SPSS Inc.'s predictive analytics applications, consisting of data mining and text analytics. The research firm found that 94% achieved a positive return on investment with an average payback period of 10.7 months on the \$342,000 they spent on average for the initial deployment.

Chances are it won't be long before text analytics software infiltrates the engineering, product marketing, compliance, customer care, and research and development departments at most major companies. Yet obstacles remain. The software isn't quite as ready out of the box as that earlier generation of structured data intelligence programs. And unlike a business intelligence dashboard sitting on a CEO's desktop, text analytics is more useful for a product engineer, which could slow adoption if senior corporate executives don't grasp the software's significance.

Perhaps most importantly, it can be difficult for startups peddling the new software to find the right group within a large organization that would derive the most value from the new product. Finding the most senior corporate executives and persuading them of the benefits of text analytics is the toughest task at this point, say the mostly startup companies selling the software. But as the benefits of text analytics filter back into the marketplace, watch for more and more companies to adopt the new technology.

Other Whirlpools

Whirlpool isn't the only corporation using text analytics to improve customer service and fix problems before warranty costs spiral into the hundreds of millions of dollars. The forklift division of Cleveland-based NACCO Industries Inc., for example, has deployed the new software to reduce major warranty claims from up to two years to six months (see story, page 14).

Pharmaceutical companies such as Wyeth are using the software to mine governmental and research institute medical databases for promising drugs, and financial institutions such as Capital One Financial Corp. are combing through their unstructured data for clues about which applicants or customers are likely to default. And Dow Chemical Co. and Eastman Chemical Co. are using text analytics to originate new areas of innovation.

Many predict that an accelerating pace of text analytics adoption is imminent. "The market has turned, and now it's looking to combine structured and unstructured data for analysis," says Amanda Reed, general

partner at venture capital firm Palomar Ventures. "The hour's upon us as companies are starting to think about whether they have problems on their hands they can solve this way."

Curt Monash, president of Monash Information Services, a consulting firm focused on analytics and data management, says that the text analytics market started to explode in the fourth quarter of last year. "I think it's good technology, so it was due to explode," he says. "Analytic technology will be the center of gravity in this decade, and that is not going to change for many years."

IBM's announcement in August that it is creating a standard for unstructured information management architecture, or UIMA, is seen as a tipping point by evangelists of the new technology. The computing giant hopes UIMA will provide a standard framework that can be used to add text analytics to any application and enable interoperability among different vendors' analytics software and other enterprise software applications. The appeal of text analytics will only grow as the software begins to function more and more easily with traditional business intelligence tools made by companies such as Cognos, Hyperion and SAS. But much depends on how well this new software ties in with what's already in the marketplace.

Analytics' evolution

Text analytics evolved from enterprise search companies such as Autonomy Corp. and Verity Corp., which pioneered unstructured data search technologies. "These guys paved the way, but like a lot of early products, people believe the capabilities of those systems are the capabilities of all the systems," says Palomar's Reed. "But Attensity is a real next-generation application."

Reed, whose firm invested in Attensity two years ago, says that the key difference is that Attensity has built applications targeting different industries, such as hospitality and manufacturing, which can be easily used by business people rather than Ph.D.s in the back room. David Bean, chief technology officer and co-founder of Attensity, explains that the company begins by diagramming sentences in order to determine its contents and context. After understanding the relationships between the key words, it places the text in a database that can then be analyzed.

Bean's work on text analytics goes back to 1990 while he was working on a project with five peers at the University of Utah. He and others at other startups began pushing in various directions to make free text found in customer call notes or service reports look like structured data.

Competitors such as ClearForest, an Israeli company backed by venture capital firm Greylock Partners, have developed specific text analytic applications for the chemicals, financial, publishing and other industries. It's working with Accenture Ltd.'s R&D group to support its effort to develop a prototype that provides information about strategies of large corporations.

ClearForest's ClearTags software mines documents on the Web, blogs and newsfeeds to identify quotes that can then be extracted and loaded into a database, which clients can search. That information in turn can be packaged as an application for technology companies to gain unique insight into what rival corporations' chief technology officers consider their next great strategic undertaking.

Stein Technologies Inc. has developed a text analytics product for the publishing industry, while Temes SA has created a niche using text analytics to help companies track their reputations. They are all attempting to create more user-friendly products than the first generation of unstructured data analysis software, which was more of a platform than an application, according to Palomar's Reed.

"Text analytics allows you to take repositories of data and allows you to understand what's in there," adds Troy Pearsall, vice president of technology at In-Q-Tel Inc., a VC firm funded by the Central Intelligence Agency. "It's a problem for all of us as the information explosion continues."

Governmental agencies such as the Federal Bureau of Investigation and CIA were the earliest adapters of this technology, using it to automate the document analysis and tagging process in order to find the not-always-obvious connections that might exist between terrorists, arms purchases and organized crime. What In-Q-Tel likes about the technology is its ability to build relationships between the words that its software extracts.

So, for example, if all a company knows is that it is looking for information about why an engine is failing, the system would not just tell it the word "fail" is appearing near the word "engine" but also "valve" and "2,000 hours." In-Q-Tel has one of the broadest portfolios of data analysis companies in the venture capital world. The portfolio spans a range of companies, from information extraction shops such as Basis Technology Corp. to data analysis startups such as Spotfire Inc. "There is a greater understanding of the need, and the market is more receptive to companies in this space," he says.

Seth Grimes, president of Alta Plana Corp., a text analysis consulting firm, notes that pharmaceutical companies are now using text analytics to discover new drug therapies. "Clinical records and scientific studies are calling out for automatic software programs to find connections between compound genomes to discover the drugs worth testing in the labs," he says. Text analytics software is useful to Big Pharma because it can help companies discover relationships between compounds they may not have expected. And it can be done quickly because it's automated.

But other industries could well adapt the software, Grimes adds. "The only requirement is having a large number of text documents and a need to analyze them, understand them and find links to business goals," he says. "There are many places where this technology can be applied."

One of Attensity's customers in the hospitality industry, for example, uses the software to cull intelligence from the free text found on customer feedback forms left behind in hotel rooms. Whatever the complaints, addressing them swiftly is a way to improve service that only free text on a customer service card will reveal.

Text analytics can be put to ever greater use, its enthusiasts say, by tapping into what customers want and then delivering it to them immediately. Whirlpool's Welke, for example, is now deploying the software to help his company's call center staff and service technicians have the right answers to customer questions at their fingertips.

Case in point: When Whirlpool recently released a new refrigerator based on decades of experience and reams of customer research, it expected few complaints. Moreover, only a handful of customers called in to say how difficult it was to fit a carton of milk into the refrigerator's side door. Previously, Whirlpool would not have noticed anyone complaining about the accessibility of the side door. There had been one refrigeration expert that call center operators could go to for help, and unless that expert spotted the complaints, they would likely have gone undetected.

Now, those complaints about the side door of the refrigerator have allowed the company to quickly identify the problem, decide upon what design change should be taken and then modify the interior design of the refrigerator in future models.

"Are you taking advantage of the millions of touches you have every year with your customer?" Welke, now vice president of operations in North America, asks rhetorically. "This is much more of a knowledge opportunity. Once I got beyond the 'I need this for early warning,' it's 'What can I learn from the consumer?'"