UpClose: Janine Toole

Creating order from chaos

by Heather Finley

Dr. Janine Toole is at an exciting point in her career. As vice-president of product development at Axonwave Software (*www.axonwave.com*), she is directly involved in the development and implementation of natural language processing (which is the engineering of systems that process or analyze written or spoken natural language) in industrial settings. And that puts her on the leading edge of business intelligence.

Imagine being able to "review" tens of, or hundreds of thousands of United States Federal Aviation Administration or Canadian Department of Transport incident reports in minutes. Then, at the end of your review, generate a bar chart that gives you an instant focus on your company's problem products. Axonwave's software currently does that for Boeing, and they're in talks with other companies who are equally interested in tracking product problems before consumers get hurt.

One example, a search on 30,000 documents relating to a child car seat manufacturer, returns 18 incident reports. Within those, the software finds reports of injuries apparently unrelated to each other, and ranks them by severity and timing so the manufacturer knows whether or not there has been an increase. The complete search and graphic output takes less than ten minutes.

Information Highways: Trace your career path.

Janine Toole: It all started in Australia, where I did a BA in linguistics. My interest was languages; Japanese, French, and linguistics itself. When I graduated I went to work for Shell Australia in their computer department. But computing wasn't quite "it," so I went back and did an MA at Monash University. Then I came to Canada to do a PhD at SFU [Simon Fraser University], keeping the focus on the balance between computing and linguistics. After I graduated, I worked in the natural language laboratory with [Axonwave founder] Fred Popowich and Paul McFetridge [chair, department of linguistics, SFU], who were faculty members at SFU. By the time I finished the post-doc, Gavagai — now Axonwave — had been going for a couple of months, so I joined Gavagai, which is actually a spin-off company from SFU.

I really like working here because, again, it's that balance between computers and language and linguistics. It's a really exciting time for the technology. Ten years ago, all this technology was still in the universities. Now I think we're going to see more and more of them coming to the marketplace over the next couple of years.

IH: What's the origin of the name Gavagai?

JT: W.V. Quine was a philosopher of language. One of the stories that he used to tell was this: say you went to a foreign country and you didn't speak their language and the people there didn't speak your language. Your host points at a rabbit running by and says "Gavagai." How do you know what he's saying? Is he saying, "rabbit?" Is he saying, "dinner?" Is he saying, "run?" You have no idea what he's actually saying. The point was that language is very indeterminate. Even when you share a language, can you really be sure that what you're saying and what the person thinks they're hearing has the same meaning? It's about the difficulties of interpreting language, and that's what we're trying to do. We chose it as our company name — which we recently changed [to Axonwave].

IH: Describe Axonwave's content intelligence system — what does it do, what are its benefits, who is it targeted to?

JT: As a company, we've decided to focus on the area of product safety. We want to help people identify early if they might be having problems with their products. For example: one of our customers, Boeing, receives over 20,000 incident reports per year from the FAA that have been submitted by pilots, mechanics, stewardesses — any time they've witnessed any sort of incident they think is safety-related.

Boeing and every other manufacturer in the aviation industry wants to know what's going on in those reports so they know what improvements to make to their process or their products. Trouble is, there are so many reports that [Boeing] says, if we get a human to analyze it the way that we want to analyze it, we can only analyze about 200 reports per year because we need to do such in-depth analysis. The earlier they can detect a problem, the better. Our software gives them the ability to focus in on the documents that are perhaps indicative of something they should be aware of. What we do is help them identify the reports that they should be looking at and help them identify if they have any problems with their products.

IH: What other kinds of applications are out there for your software?

JT: One of the challenges of any early stage company, particularly NLP-based companies, is finding a market to focus on and being able to describe the problem you solve, not just how cool your technology is. That's why

we're focusing on the product-monitoring area. But there are definitely numerous other applications for our software. For example, we've had some discussions with a Canadian law firm that is interested in identifying confidential information in a document. They want to be able to identify anything like a name, address or any information that could identify who an individual is. A parallel use of that is in the government, where they need to respond to freedom of information requests but they need to obscure the confidential information.

IH: One of the features of Axonwave is being able to visualize the result. It's not just getting a list of things, it's being able to get that in five minutes and being able to get a very quick visual about what the trend is.

JT: Exactly, essentially, it's to summarize. Business intelligence works on structured data — the formulated data you have in databases. We see ourselves as the exact parallel in the unstructured world, where we allow you to analyze and visualize the unstructured data that's underneath.

The BI world has been growing hugely in the last couple of years while all the other software companies are struggling. They're taking the data which people have been storing in databases and data warehouses; they're actually doing the next step of making that data useful for business decisions. That's the exact parallel of what we've been trying to do on the text or content or unstructured information side. We know we've been storing it for years: we've got it in document management systems, content management systems, and it's on the Web. We've got to take that next step and make it useful for business decisions, otherwise there's no reason to keep it. That is why that visualization part is so valuable; it's taking the massive data and converting it to something from which we can ask the next question or make a decision, or at least move forward in solving our business problem.

It [could be] charts but it could be a table of numbers; there are many different ways we could have presented the information. A chart is just something that humans find easy to interpret. Again, it's one piece of the whole problem, which is going from that unstructured data and converting it to a format that lets you make a decision.

IH: Are you looking at languages other than English at this point?

JT: We have prototypes for French and Spanish because they are the obvious next step. At the moment we're not going beyond the prototype stage because there hasn't been any sort of market need that we've seen. The prototypes we did were really more a proof of concept to show "yes we can." We've developed the software so it's quite modular. It's just a matter of taking out the English module — like the grammar, the lexicon and things like that — and plugging in French or Spanish or whatever other language is of interest.

IH: Where is NLP as an industry headed?

JT: I think it's a really interesting time for NLP because there's a bunch of companies out there. I think the ones that are going to be successful are those that have moved beyond talking about the "coolness" of the technology to the business problem they're solving. Rather than lots and lots of NLP companies, there will be companies that are solving particular business problems like we are with product monitoring. That means that NLP is finally coming of age, as it were. People are talking less about the technology and more about the business applications. ih

Heather Finley is a freelance writer based in Toronto. She can be reached at heather@illuminarts.com.