Natural Language Engineering **11** (4): 435–438. © 2005 Cambridge University Press
doi:10.1017/S1351324905003979 Printed in the United Kingdom

Industry Watch

ROBERT DALE
Centre for Language Technology, Macquarie University, Sydney, Australia

(*Received 4 October 2005*)

Suppose you're a corporate vice president at a well-known international software company, and you want to check on the visibility of one of your leading researchers in the outside world. You're sitting at your desk, so the most obvious thing to do is to enter their name into a search engine. If the well-known international software company happened to be Microsoft, and if the leading researcher happened to be Microsoft's Susan Dumais, and if the search engine you decided to use happened to be Google, you might be surprised to find that the sponsored link that comes atop the search results is actually from Google itself, exhorting you to 'Work on NLP at Google', and alerting you to the fact that 'Google is hiring experts in statistical language processing'.

Well, the sponsored ad probably wasn't what enticed a particular corporate vice president, Kai-Fu Lee, to jump ship at Microsoft in July of this year; a rumoured $10 million might have more to do with that.[1] But it's not impossible that there's a causal link in the other direction: perhaps the sponsored ads are a little Google-swipe at Microsoft in light of the recent legal action over Lee's departure from Microsoft. Just in case you have not been watching the news: on 19th July, Google announced that it had hired Lee to head its Chinese research and development team. Microsoft immediately filed a lawsuit against Lee and Google alleging breach of Microsoft's employee confidentiality and non-compete agreement. It all comes down to what Lee worked on, and has inside knowledge of, at Microsoft, and how that knowledge might impact on his work at Google. Lee says he had only worked on natural language technology with Microsoft, but Microsoft say he worked on search issues in direct competition with Google. I'm just a little worried that someone will use this as evidence that search has nothing to do with natural language technology. As I write this in late September, Judge Steven Gonzalez of the Superior Court of Washington State in King County has ruled that Lee can begin work for Google setting up the R&D center in China while a lawsuit over his hiring by Google awaits a trial that will start in January 2006.

If you've got a pre-Web memory, you'll be aware that Kai-Fu Lee completed his doctoral dissertation in 1988 at CMU on the Sphinx large vocabulary speech

[1] The apparent linking of Susan Dumais's name to jobs at Google was noted in various blogs and news items around the end of August 2005; I haven't been able to find any earlier mentions of the phenomenon, so I don't know when it first appeared, and it may be gone by the time you read this.

https://doi.org/10.1017/S1351324905003979 Published online by Cambridge University Press

recognizer, an application that's still going strong, before going on to work on speech technologies at Apple. He moved to Microsoft, after a stint at SGI, in 1998. Speech technology is at the core of another recent, if lower-key, legal wrangle. Yahoo wants to integrate speech technology into its web services, and has been working with Nuance on this. But in late September, Larry Heck, Nuance's VP of R&D, and a dozen Nuance engineers left to join Yahoo. Nuance claims that the technology they were working on for Yahoo was 75% complete, leading Nuance to conclude that Yahoo is attempting to swipe its technology after it had done all the hard work. At the time of writing, it was reported that Nuance planned to sue Yahoo.

But just who is Nuance anyway, these days? It's all getting a bit confusing: not very long ago there were really two major players in the speech world, those being Nuance and SpeechWorks. Of course there were other vendors of speech engines out there, but between them the two biggies had locked up the bulk of the market for telephony-based speech recognition systems, providing the underlying technology that drove applications developed by an army of third-party developers. But market share doesn't mean financial success, especially in the speech business. It was widely acknowledged that both companies were burning money faster than they could make it, so the real question was who would have to sell out first. SpeechWork's strategic alignments with Microsoft in 2002, including support for Microsoft's SALT (Speech Application Language Tags) technology, raised some suspicions that we'd see a Redmond takeover. But in April 2003, SpeechWorks was acquired by ScanSoft.

Now, ScanSoft, as you may recall, was originally a wholly-owned subsidiary of Xerox. In March 1999 it was acquired by Visioneer, the developer of the PaperPort range of scanning devices. Visioneer sold off its hardware arm, and retained the ScanSoft name for the software side of the business. For a while the company was viewed as a home for orphaned digital imaging products: in 2000, it acquired Caere, another company that produced optical character recognition software. Then, at the end of 2001, ScanSoft acquired all of the assets of Lernout & Hauspie's speech and language technology business, including the well known Dragon Naturally Speaking range. This paved the way for more acquisitions in the speech world. Emulating L&H's earlier role as a gobbler-up of speech and language technology companies, in early 2003 ScanSoft completed its acquisition of Philips' speech processing, telephony and voice control business units, followed by the SpeechWorks acquisition mentioned above. In late 2004, it signed agreements to acquire Phonetic Systems, ART Advanced Recognition Technologies, and Rhetorical Systems.

I've probably missed some strategic takeovers along the way (blame a low-recall named entity recognizer), but the most stunning event is surely the May 2005 announcement by ScanSoft of a merger with Nuance. With a combined 75% of the market, this is just the kind of thing to raise antitrust concerns. When they were competitors, each of the companies spent years marketing themselves to customers as the only viable players in the speech world, arguing that IBM and Microsoft were bit players in this arena. The merged ScanSoft/Nuance had to convince the US Federal Trade Commission that it would not be an effective monopoly, on the grounds that IBM and Microsoft are real competitors after all. The FTC have given the go ahead, and the merger was completed in mid-September 2005. All this

https://doi.org/10.1017/S1351324905003979 Published online by Cambridge University Press

reminds me of the strange feeling I had upon discovering that the company that made my breakfast cereal was owned by the same conglomerate as the one that made the bread I had just toasted, as well as the company that made the jam I was going to spread on the toast, not to mention the cat food that Gizmo, our cat, was in the process of eating.

Which in a roundabout way brings us to named entity recognition, and other related intelligent text processing tools. It's great to see the stuff we talk about at conferences finding its way into publicly accessible products. For a while there, the best (and often most amusing) way to see how you, your friends or your colleagues fared on the net was to type the relevant name into `www.googlism.com`; fun, but not always maximally informative. But real cross-document coreference is alive and kicking on a web site near you: check out `www.zoominfo.com`. ZoomInfo does information extraction by crawling the web. It looks like it uses information about affiliations and places of work to cluster references to individuals, allowing it to distinguish different people with the same name. Although, inevitably, it's not perfect, it does a pretty impressive job.

Perhaps we'll begin to see more and more applications like this, as a consequence of another big language technology news item this year: IBM announced its intention to make available via open source its Unstructured Information Management Architecture, or UIMA for short. UIMA defines a common, standard interface that enables intelligent text processing components (or, as they are increasingly referred to these days, text analytics modules) from multiple vendors to work together. This is potentially serious stuff, and something of a breakthrough for the relatively fragmented text processing world: simultaneous with the announcement, more than 15 vendors, including Attensity, ClearForest, Cognos, Endeca, Factiva, Kana, Inquira, iPhrase, Inxight, nStein, QL2, SAS, Schemalogic, Semagix, SPSS Inc. and Temis, pledged to support UIMA in their products – that's pretty much everybody who's anybody in the business. The move toward a plug-and-play architecture is good news for developers and users of these technologies, since in theory it promotes a mix-and-match approach to building text mining applications, rather than requiring an approach where everything is sourced from a single vendor. In this regard, UIMA might have a similar impact to that which has been claimed for VoiceXML in the speech industry. The Software Development Kit that IBM is making available provides a collection of tools and APIs including what is referred to as 'semantic search'; find out more at `www.alphaworks.ibm.com/tech/uima`.

Meanwhile, if all you really want to do is ask simple questions about people, you might be aware that Google has joined Ask Jeeves and MSN Search in offering a basic question answering capability: the answer to a question comes up as a specially marked item at the top of the search results. I got correct answers to *Who is the president of the USA?* and *Who is the prime minister of Australia?*. On the other hand, although *Who is Steve Ballmer?* tells me that he's the Chief Executive Officer of Microsoft, asking *Who is the the Chief Executive Officer of Microsoft?* doesn't return a specially-marked answer (although of course the returned web

https://doi.org/10.1017/S1351324905003979 Published online by Cambridge University Press

pages provide an answer to the question). *Who is Kai-Fu Lee?*, on the other hand, returns the answer '(Former) Corporate Vice President, Natural Interactive Services Division at Microsoft'.

I expect many people who, like me, teach a course on spoken language dialog systems will have a collection of standard scenarios they use to communicate the potential for the technology to students. Along with talking to your microwave, usually right at the top of the list is talking to your TV or video recorder. That's passé, it seems: you can do it now, thanks to One Voice's Media Center Communicator, for US$150; see `www.onev.com/mcc/` for a demo. Time to think of a more elaborate device interaction to keep the students interested.

https://doi.org/10.1017/S1351324905003979 Published online by Cambridge University Press