

Expert System: taking a poke at Google's eye

Mike Davis and Madan Sheina

Expert System, a company that started out by making language software for Microsoft's spell check software in the late 1980s, is one of several vendors poking their fingers in the eye of Google's keyword search by applying semantic search technology to disambiguate search queries and web text in order to increase the precision and relevancy of results.

Keywords running out of productivity gas

Currently keyword search remains the most popular search technique for users on the public Web and corporate intranets. But many believe its time is up because consumers and business users no longer want to see 30,000 hits on a search and then wade through a list of loosely related keyword results to find relevant documents.

This where a new breed of so-called semantic technologies comes into the frame. Unlike ranking algorithms such as Google's PageRank for predicting relevancy, semantic search dips into the meaning in language to produce highly relevant search results.

For example, Expert System provides its own semantic search platform – branded as Cogito (Latin for "I think"), and is provided as a fully-hosted service worldwide to offer businesses 'better search'.

Semantic search

The Cogito semantic engine is designed around the principles of human comprehension to allow content to be understood in the way in which the author intended it to be. This is something that keyword search ignores. For example, a Google search for the word 'jaguar' would pull up content around the animal and the car. Semantic search would look not only at the keyword but also other words around it like 'jungle' or 'saloon' to separate the two meanings.

Semantic search is just one of several search techniques that are being forwarded as better and more precise alternatives to keyword search. Others include heuristics and ontology, linguistics and text mining, and statistical. However, Expert System claims that these approaches fall short, addressing only the morphological and grammatical aspects of analysis.

What semantic search does is look at sentence logic (how words in a sentence relate to one another) and semantic analysis (understanding the context of keywords – referred to as 'word sense disambiguation' in semantic parlance). When a term is ambiguous, meaning it can have several meanings (for example, bark), it needs

some kind of semantic analysis of the other words that wrap around it to give it its true meaning and context.

Other search engines often hit a brick wall when it comes to deep analysis. For example, when a heuristically-driven search engine sees two adjectives in a sentence it usually washes them out and scores the sentence as neutral because it has no understanding of where the two separate adjectives are pointing.

Semantic networks – the secret sauce

Expert System believes it can go the extra mile because it has a semantic network – a lexical database that provides a knowledge representation of word definitions and their relationships. In essence it has poured Webster's dictionary into an in-memory database – comprising 350,000 words and 2.8 million relationships.

Importantly, Expert System's semantic network also focuses on common words. That's different from most ontological approaches that concern themselves with wrapping meaning and context around specialised (often scientific) content and skip common words that comprise 90% of all content.

Riding the Web 3.0 wave

Expert System isn't the only company eyeing the semantic web (currently dubbed 'Web 3.0'). Other semantic start-ups include Powerset, Yedda, Trovix and Hakia. Awareness of semantic search rose this summer when Microsoft picked up San Francisco-based Powerset. Interestingly, Microsoft followed that up with its second semantic buy – Zoomix, a data quality provider that has baked semantic self-matching methods into its software. These moves are important. Up to now semantic search has been a market where there has been a lot of theoretical hype but little real substance or proof that it works better than current search technology.

Semantic networks are tricky to build and not all are equal. However, it's unlikely that Expert System and other semantic technologies will ever be able to provide 100% precision in their analysis and results. Moreover there are still question marks over potentially sticky performance issues with semantic searches that eat up more processing cycles.

A short report on Expert System on Cogito will be published on the Knowledge Center in the near future, and Ovum is currently partnering in a major, two-year EU research project on Semantic Technologies for the Enterprise (<http://value-it.isoco.net/>)

Ovum does not endorse companies or their products. Ovum operates under an Independence Charter. For full details please see www.ovum.com/about/charter.asp.

For full details of Ovum's citation policy, see www.ovum.com/media/citation.asp.

Whilst every care is taken to ensure the accuracy of the information contained in this material, the facts, estimates and opinions stated are based on information and sources which, while we believe them to be reliable, are not guaranteed. In particular, it should not be relied upon as the sole source of reference in relation to the subject matter. No liability can be accepted by Ovum Europe Limited, its directors or employees for any loss occasioned to any person or entity acting or failing to act as a result of anything contained in or omitted from the content of this material, or our conclusions as stated. The findings are Ovum's current opinions; they are subject to change without notice. Ovum has no obligation to update or amend the research or to let anyone know if our opinions change materially.