

Wyszukiwarka domyśli się o co chodzi

12 października 2007

Badacze z francuskiego oddziału firmy Xerox opracowali FactSpotter – narzędzie wyszukiwawcze analizujące gramatyczną stronę poszukiwanego tekstu.

Dzięki technologii użytej w FactSpotterze, wyszukiwarka będzie mogła rozpoznawać znaczenie słów wieloznacznych, np. określić czy dane słowo zostało użyte jako rzeczownik czy czasownik. To zaś pozwoli na bardziej precyzyjny dobór wyników generowanych przez wyszukiwarę.

W czasie prac nad FactSpotterem naukowcy stworzyli własny metajęzyk opisujący gramatykę różnych języków, na razie kilku: angielskiego, francuskiego, niemieckiego, niderlandzkiego, włoskiego, portugalskiego i hiszpańskiego.

Ponadto połączony zespół Fujitsu-Xerox przygotował też na potrzeby FactSpottera opis języka japońskiego, wykazując w ten sposób, że metajęzyk Xeroxa może być użyty również względem języków wykorzystujących inny system zapisu.

FactSpotter został napisany w języku programowania C. Istnieją też moduły w Javie i Pythonie. Analizuje jedynie język pisany, ale może współdziałać z systemami rozpoznawania mowy. To zaś pozwala wykorzystać go przy np. przeszukiwaniu stacji radiowych i telewizyjnych.

Nie jest to pierwszy patent Xerox w badaniach nad wyszukiwaniem w języku naturalnym. Prace tego typu firma prowadzi od wielu lat, głównie pod kątem robotyki. Na technologii licencjonowanej przez tą firmę ma być oparta między innymi „rozumiejąca” język wyszukiwarka Powerset.

Opracowanie: Marcin Opolński

Źródło: [Dziennik Internaytów](#)