

# W internecie będziemy mogli szukać po swojemu

Grażyna Zawada

14.04.2011 aktualizacja: 2011-04-13 21:58



prof. Wiesław Lubaszewski z Katedry Lingwistyki Komputerowej UJ i Katedry Informatyki AGH

☒ Fot. Jakub Ociepa / Agencja Gazeta

**Lingwiści komputerowi z Krakowa opracowali inteligentny system wyszukiwania informacji internetowych. Technologia wejdzie w skład Polskiej Platformy Bezpieczeństwa Wewnętrznego.**



☒

Fot. Jakub Ociepa / Agencja Gazeta

prof. Wiesław Lubaszewski z Katedry Lingwistyki Komputerowej UJ i Katedry Informatyki AGH

- Dzień dobry, szukam fajnych dzinsów, nie za drogie, ale dobrze uszytych - takie objaśnianie naszych potrzeb mogło jak dotąd odbywać się jedynie w rozmowie z człowiekiem. Gdyby zapytać o to samo zwykłą wyszukiwarkę internetową, prawdopodobnie pokazałaby nam dzinsy statystycznie najczęściej wybierane przez innych klientów. Czy kiedykolwiek komputery rozumieją, o co dokładnie chodzi człowiekowi? Jest szansa. Zespół naukowców z Akademii Górniczo-Hutniczej i Uniwersytetu Jagiellońskiego opracował technologię zindywidualizowanego wyszukiwania informacji w internecie.

- Maszyna znajdzie dla nas wszystko, jeśli prawidłowo zdefiniujemy to, czego szukamy. Do dyspozycji mamy jednak wieloznaczeniowe słowa. Komputer ich nie zrozumie, ale może je dla nas znaleźć. Problem w tym, jak mu przedstawić znaczenie słów, żeby wybierał tylko treści dla nas istotne - mówi prof. Wiesław Lubaszewski z Katedry Lingwistyki Komputerowej UJ oraz z Katedry Informatyki AGH, koordynator zespołu, który zbudował system o nazwie Monitoring Polskojęzycznego Internetu (MPI).

Naukowcy wyszli z założenia, że skoro wyszukiwanie informacji jest subiektywne, to należy użytkownikowi udostępnić narzędzie, które pozwoli mu szukać "po swojemu". - Dopiero wtedy selekcja informacji będzie naprawdę satysfakcjonująca - przekonują twórcy MPI. Komercyjna wyszukiwarka stosuje model statystyczny i zwraca wiele tekstów nie na temat lub gubi te, które zawierają informacje opisowe. - MPI uwzględnia znaczenie słów w pożądanym przez szukającego kontekście. Skojarzy frazy "wypił herbatę" i "rozkoszował się każdym łykiem" albo znajdzie oferty samochodu, gdzie słowo "tani" oznacza 5 tys. zł, a nie 50 - mówi Lubaszewski.

Na razie prototypowa wersja systemu została dopasowana do potrzeb policji. W fazie testów zadaniem wyszukiwarki było wyselekcjonowanie ogłoszeń dotyczących handlu organami do przeszczepu. Tylko na jednym średniej wielkości portalu ogłoszeniowym w ciągu kilku godzin nocnych pojawiły się aż 253 ogłoszenia kupna i sprzedaży organów do transplantacji. Wyszukiwarka odnalazła je w ciągu kilkunastu sekund. Taki sam mechanizm można zastosować do śledzenia w sieci śladów wszelkiego typu przestępstw. - Znalezione treści rozpracowują przeszkoleni funkcjonariusze, którzy prowadzą dalszą pracę dochodzeniowo-śledczą - mówi prof. Lubaszewski.

Z wyszukiwarki mogłoby także skorzystać wiele osób zajmujących się monitorowaniem treści internetowych pod kątem badania opinii społecznej lub plagiatów. Aby jednak MPI był dostępny na zasadach komercyjnych, potrzeba kilku lat badań i stałych udoskonaleń. - Moi młodszy koledzy nie mogą sobie pozwolić na poświęcenie czasu MPI, bo tworzenie technologii praktycznie nie liczy się do dorobku akademickiego. Gdyby zmienić obowiązujące przepisy, więcej nowoczesnych technologii opracowywanych w Polsce wyszłoby poza stadium prototypu - mówi prof. Lubaszewski.

Więcej...

[http://krakow.gazeta.pl/krakow/1,44425,9431466,W\\_internecie\\_bedziemy\\_mogli\\_szukac\\_po\\_swojemu.html#ixzz1JsZKMz51](http://krakow.gazeta.pl/krakow/1,44425,9431466,W_internecie_bedziemy_mogli_szukac_po_swojemu.html#ixzz1JsZKMz51)