



# Igła w stogu siana

Ilość informacji dostępnych poprzez World Wide Web jest przytłaczająca. Aby nie stracić orientacji podczas wyszukiwania w internetowych zasobach, najlepiej posłużyć się specjalnymi wyszukiwarkami.

**S**pecjaliści od strategii reklamowych znaleźli dla siebie nowe medium: zewsząd słycać nawoływania zachęcające do zagłębienia się w cyberprzestrzeni. Te zachęty trafiają na podatny grunt. Coraz więcej osób chce się przekonać, czy surfowanie po Internecie jest rzeczywiście wciągającym zajęciem. Jednak nieopierzonym internautom często brakuje jakiegokolwiek orientacji w przebogatych zasobach stron World Wide Web. Jeśli ktoś szuka określonych informacji i nie wie jak do nich dotrzeć, może się bardzo szybko zniechęcić do regularnego eksplorowania Internetu.

Wszystko jedno czy szuka się adresu e-mailowego, artykułu z grupy tematycznej (newsgroup), programu shareware'owego, czy też specjalnej oferty jakiejś firmy: bez wsparcia w postaci odpowiednich narzędzi podczas wyszukiwania jest się na straconej pozycji.

Na szczęście pomoc jest w zasięgu kliknięcia myszką. W wytropieniu pożądaných informacji oprócz niezliczonych katalogów zawierających odnośniki (linki) do dokumentów Webu, przydatne są przede wszystkim tzw. wyszukiwarki. Można je określić jako ogromne online'owe bazy danych, które na podsta-

wie słów kluczowych wyszukują strony w WWW. Gdy wpisze w odpowiednie pole wyszukiwarki słowa identyfikujące interesujący nas temat, w wyniku wyszukiwania otrzymamy listę adresów sieciowych (URL) odnoszących się do znalezionych dokumentów.

Istnieją też podobne usługi przydatne w wyszukiwaniu adresów e-mailowych, oprogramowania i artykułów, pojawiających się w grupach tematycznych. Poza aktualnością i kompletnością bazy danych o jakości wyszukiwarki decyduje przede wszystkim elastyczność systemu formułowania pytań wyszukiwawczych.

Nie znaczy to jednak, że wyszukiwarki w prosty sposób rozwiązują problem wyszukiwania informacji w World Wide Web. Skuteczność wyszukiwania wymaga pewnej wprawy. Gdy ktoś przykładowo chce zasięgnąć informacji na temat katastrofy promu kosmicznego Challenger i wpisze w polu wyszukiwarki AltaVista słowo „Challenger”, otrzyma całą masę adresów spełniających zadane kryteria. Wśród nich znajdują się również informacje o Challenger District w amerykańskim stanie Minneapolis czy przedsiębiorstwie geologicznym Challenger Minerals. Oczywiście nie takich odpowiedzi spodziewał się użytkownik zainteresowany nieudanymi lotami w kosmos.

Dlatego serwery umożliwiające wyszukiwanie oferują z reguły bardziej skuteczne metody. Za pomocą operatorów AND, OR i NOT można sprecyzować pytanie i dzięki temu zredukować listę odpowiedzi na dany temat. Jeżeli, korzystając z wyszukiwarki AltaVista, do opisanego wcześniej pytania dodamy wyraz „Shuttle”, wówczas serwer „wygeneruje” znacznie mniej adresów.

CHIP przyjrzał się najważniejszym wyszukiwarkom i udostępnianym przez nie opcjom. Okazało się przy tym, że polowanie na informacje, mimo wspomaganie ze strony różnych wyszukiwarek i tematycznie uporządkowanych katalogów, wciąż jest zajęciem dość żmudnym. Kto żąda precyzyjnych wyników wyszukiwania (trafień), musi za każdym razem „rozprawić się” ze specyficznymi mechanizmami zadawania pytań.

Z pewnością wielu użytkowników wołałoby, żeby obsługa wyszukiwarek była bardziej elastyczna i w niektórych miejscach zorganizowana w łatwiejszy sposób. Poza tym brakuje usług, które koncentrowałyby się na poszczególnych dziedzinach. Takie wyszukiwarki byłyby bardziej przejrzyste i być może także nieco szybsze niż większość aktualnie działających „wszechwiedzących” maszynek do wyszukiwania.



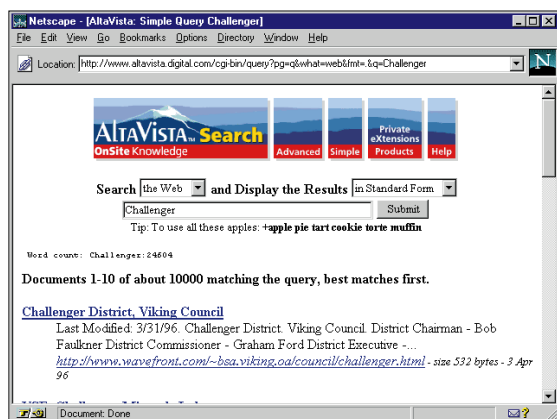
### AltaVista

<http://www.altavista.digital.com>



Jeśli wierzyć statystyk, AltaVista jest jednym z największych indeksów w Internecie. Obejmuje ponad 30 milionów stron zgromadzonych na około 275 000 serwerów. Cechą szczególną tej popularnej przeglądarki jest to, iż może służyć nie tylko do przeszukiwania WWW, lecz także do wyznajdowania artykułów publikowanych za pośrednictwem grup tematycznych. AltaVista udostępnia użytkownikom prawie trzy miliony wiadomości sieciowych pochodzących z ponad 14 000 grup.

Użytkownik może określić trzy sposoby (od kompaktowego po szczegółowy) dokumentowania przez serwer wyników wyszukiwania. Zaawansowane wyszukiwanie można rozpocząć wybierając link *Advanced Search*. Wówczas uzyskuje się



**AltaVista: dzięki zmyślnej procedurze wyszukiwania i krótkiemu czasowi odpowiedzi wyszukiwarka ta stanowi alternatywę dla Yahoo czy Lycosa**

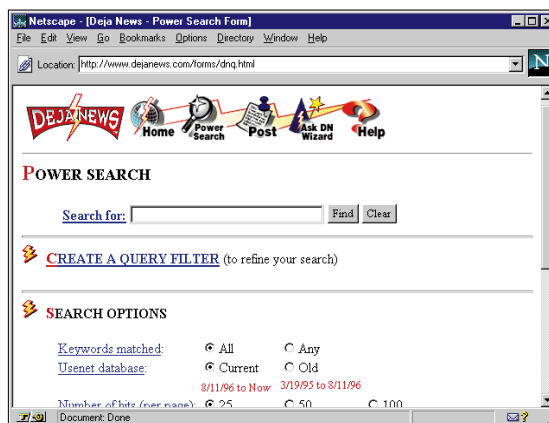
możliwość użycia operatorów AND, OR, NOT i NEAR. Ich zastosowanie wyjaśnia szczegółowo strona pomocy.

### Deja News

<http://www.dejanews.com>



Mimo iż istnieje wiele wyszukiwarek, które oprócz Webu przeszukują także grupy tematyczne Usenetu, w szczególnych przypadkach warto odwołać się do fachowca: Deja News jest przeznaczony specjalnie do „przechowywania” grup tematycznych i umożliwia dostęp do wciąż rozrastającego się archiwum wiadomości, które obecnie sięga wstecz do marca 1995 roku. Ambitnym celem operatorów Deja News jest przygotowanie do wyszukiwania wszystkich kiedykolwiek przesłanych przez



**Deja News: specjalista od wiadomości sieciowych przekopuje się przez ponad 20 000 grup tematycznych Usenetu**

Internet wiadomości. Gdyby powiedział się ten zamiar, w archiwum zostałyby zgromadzone dane od 1979 roku.

Najprostszy i najszybszy sposób wyszukiwania (*Quick Search*) można rozpocząć korzystając ze standardowego formularza dostępnego bezpośrednio na stronie głównej. Pytanie może zawierać operatory logiczne i różne symbole, które pozwalają na przykład upewnić się, iż wyszukiwanie dotyczy tylko wyznaczonej grupy tematycznej bądź określonego autora.

Wyczerpujące angielskojęzyczne wyjaśnienia, wraz z konkretnymi przykładami, pozwalają nie wprawionym użytkownikom zapoznać się ze sposobem definiowania pytań. Jeśli ktoś pragnie skorzystać z jeszcze bardziej zaawansowanych opcji, może sięgnąć po *Power Search*. Wówczas możliwe jest między innymi kontrolowanie formatu wyświetlania, czasu utworzenia wiadomości oraz sortowanie rezultatów wyszukiwania.

### Excite

<http://www.excite.com>



Korzystając z tego gigantycznego indeksu stron WWW można wyszukiwać spośród ponad 50 milionów dokumentów – od A do Z. Dzięki specjalnej procedurze wyszukiwania (ICE, Intelligent Concept Extraction) system ten generuje odpowiedzi w zaskakująco krótkim czasie.

Excite znacznie ułatwia użytkownikowi wprowadzanie słów wyszukiwanych: system radzi sobie z pytaniami sformułowanymi w języku naturalnym.

Nawet wpisanie wyrażenia w języku polskim, mimo pewnych przekłamań, często prowadzi do sensownych wyników.

Precyzję wyszukiwania można też kontrolować poprzez operatory: wystarczy tylko kliknąć odnośnik *Advanced Search* bądź też wpisać bezpośrednio w polu wyszukiwania kilka słów oddzielonych operatorami AND, OR albo AND NOT. Operatory muszą być wpisane wielkimi literami. W obrębie wyrazów składających

się na pytanie użytkownik może utworzyć grupę. Osiąga się to poprzez ograniczenie nawiasami słów tworzących logiczną całość, na przykład „surf AND ((North Sea) OR (Baltic Sea))”.

Dodatkową funkcją Excite’a są różnorodne możliwości dostępu do amerykańskich numerów telefonicznych, faksowych i katalogów adresowych.

### FTP search

<http://ftpsearch.ntnu.no/ftpsearch>

#### FTP search v3.3

Do wyszukiwania software’u, który można za darmo ściągnąć z serwerów, znakomicie nadaje się FTP search. Zadanie pytania w tym systemie powoduje wyświetlenie listy adresów serwerów FTP, wraz z takimi dodatkowymi informacjami takimi, rozmiar pliku, data jego utworzenia i ścieżka do pliku na danym serwerze.

Użytkownik może zawęzić obszar wyszukiwania do określonej domeny, na przykład do wybranych krajów lub instytucji. Można też określić maksymalną liczbę wyszukanych plików. W przypad-



**FTP search: wytrawny tropiciel plików znajdujących się na serwerach FTP**

ku bardziej szczegółowego wyszukiwania interesująca jest możliwość wykorzystania języka zapytań Perl, który oferuje znaczną elastyczność, ale równocześnie wymaga pewnej wiedzy i praktyki.



### Możliwości wyszukiwarek internetowych

**CHIP**

Nazwa	Zakres wyszukiwania	Operatory i opcje wyszukiwania	Adres (URL)
AltaVista	WWW, newsgroups, e-mail	AND, OR, NOT, NEAR, przedział czasu	<a href="http://www.altavista.digital.com">http://www.altavista.digital.com</a>
Deja News	newsgroups (Usenet)	AND, OR, AND NOT, NEAR, różne operatory kontekstowe i symbole	<a href="http://www.dejanews.com">http://www.dejanews.com</a>
Excite	WWW, newsgroups	AND, OR, AND NOT	<a href="http://www.excite.com">http://www.excite.com</a>
FTP search	software z serwerów FTP	język zapytań Perl	<a href="http://ftpsearch.ntnu.no/ftpsearch">http://ftpsearch.ntnu.no/ftpsearch</a>
HotBot	WWW	AND, OR, czas, miejsce, typ mediów	<a href="http://www.hotbot.com">http://www.hotbot.com</a>
Infoseek	WWW, newsgroups, e-mail	AND, OR, NOT	<a href="http://www.infoseek.com">http://www.infoseek.com</a>
Internet Sleuth	cały Internet	wewnętrznie: AND, OR, NOT, język Perl	<a href="http://www.isleuth.com">http://www.isleuth.com</a>
Lycos	WWW	AND, OR, NOT	<a href="http://www.lycos.com">http://www.lycos.com</a>
WhoWhere	adresy e-mail	brak	<a href="http://www.whowhere.com">http://www.whowhere.com</a>
Yahoo	WWW, newsgroups, e-mail	AND, OR, przedział czasu	<a href="http://www.yahoo.com">http://www.yahoo.com</a>

#### HotBot

<http://www.hotbot.com>


Zawierający 54 miliony zaindeksowanych dokumentów HotBot należy do najbardziej wydajnych wyszukiwarek. Oprócz wyszukiwania oferuje też kilka dodatkowych funkcji. Na dość niepozornie wyglądającej stronie głównej HotBota umieszczony jest przycisk *Expert*. Po jego kliknięciu pojawia się formularz, który pozwala na ustawienie wielu opcji. Przykładowo można ograniczyć wyszukiwanie do określonego okresu i miejsca (domena, adres WWW lub obszar geograficzny).

Ciekawa jest też możliwość wyszukiwania stron WWW zawierających elementy określonego typu (*Media Type*) takie, jak Java, Audio, Shockwave czy też VRML. W praktyce funkcja ta może być przydatna na przykład dla użytkowników poszukujących aktualnych dokumentów, które wykorzystują techniki multimedialne.

#### Infoseek

<http://www.infoseek.com>


Chociaż Infoseek przeznaczony jest głównie do przeszukiwania zawartości stron WWW, to jednak oferuje on kilka innych usług, które zachęcają do szperania w Internecie. Spośród dostępnych serwisów można wybrać między innymi rubrykę poświęconą aktualnościom, uzyskać notowania kursów giełdowych i przeglądać mapy amerykańskich miast.

Wprawdzie Infoseek, podobnie jak inne wyszukiwarki, udostępnia wymyślne opcje wyszukiwania, jednak ich obsługa jest relatywnie niewygodna. Za-

miast powszechnie stosowanych operatorów użytkownik musi się mozolić z regułami wprowadzania słów, znakami szczególnymi i nawiasami. Podczas wpisywania nazwy własne i tytuły należy rozpoczynać wielką literą, a przy wymienianiu kilku nazw trzeba je oddzielać przecinkami.

Wyrazy, które w dokumencie powinny występować obok siebie, należy ująć



**Infoseek: ten wszechstronny serwer jest czymś więcej niż prostą wyszukiwarką. Tematycznie posortowany katalog zachęca do penetrowania Internetu**

w podwójny cudzysłów. Jeśli użytkownik zapomni o cudzysłowie, wtedy Infoseek zinterpretuje kolejne słowa, tak jakby były przedzielone operatorem OR. Z kolei połączenia operatorem AND definiuje się poprzez umieszczenie przed danym słowem znaku „+”. Analogicznie znak minus spełnia funkcję operatora NOT. Aby para oddzielnie wpisanych słów była podczas wyszukiwania traktowana jako nierozłączna całość, użytkownik musi wstawić pomiędzy nimi zamiast spacji łącznik; na przykład „Sushi-restaurants”.

#### Internet Sleuth

<http://www.isleuth.com>


Internet Sleuth („Internetowy tropiciel”) pozwala na przeszukiwanie 1500 katalogów, indeksów i baz danych, które użytkownik może wybrać z tematycznie uporządkowanej listy.

Z jednej strony, korzystając z listy kategorii można samemu przeszukiwać bazy danych dotyczące określonej dziedziny, co jest sensownym rozwiązaniem ze względu na olbrzymie ilości informacji. Z drugiej zaś strony, można wykorzystać opcje zapytań umożliwiające przeszukiwanie całego Internetu. Użytkownik może więc „zagonić” do pracy którąś z dziesięciu popularnych wyszukiwarek. Przeszukiwane są nie tylko indeksy stron WWW, lecz także Usenet, kolekcje shareware’u i archiwa serwisów informacyjnych takie, jak Washington Post czy CNN.

#### Lycos

<http://www.lycos.com>


Do atrakcyjności Lycosa przyczynia się przede wszystkim aktualność, prędkość działania i sposób sortowania informacji. Aby uzyskać dokładniejsze odpowiedzi przy wyszukiwaniu na tym serwerze można stosować operatory AND i OR. Odpowiedni formularz wyświetlany jest po kliknięciu linku *Custom Search*. Mimo że NOT nie pojawia się na liście opcji, to jednak operator ten jest również dostępny: przed słowem, które w kombinacji z innym powinno być



zignorowane podczas wyszukiwania, należy postawić znak minus. Wpisanie wyrażenia „czekolada -mleczna” sprawi, że odnalezione zostaną dokumenty, w których występuje słowo „czekolada”, ale nie znajduje się wyraz „mleczna”. Jeśli za wyrazem zostanie postawiona kropka, jak na przykład „bank.”, wtedy Lycos znajdzie tylko te dokumenty, które zawierają dokładnie słowo „bank”, a nie na przykład „banki” albo „bankier”.

Ciekawą możliwością jest wyszukiwanie z góry określonej liczby trafień, z której użytkownik może skorzystać na przykład wtedy, gdy chce dotrzeć do dokumentów zawierających słowa „Pasadena” i „USA”, a nie ma pewności, czy „Pasadena” pisze się przez jedno czy przez dwa „s”. Lycos roz-

Wówczas zostaną wyświetlone adresy stron WWW, które zawierają tylko dwa z trzech podanych w pytaniu terminów. Powiązanie wyrazów wyłącznie operatorem OR (ze względu na częste występowanie słowa „USA”) doprowadziło do wyszukania mnóstwa kompletnie nieprzydatnych dokumentów.

### WhoWhere

<http://www.whowhere.com>



Każdy, kto porusza się po sieci Internet bądź jako aktywny uczestnik dyskusji w grupach tematycznych, bądź też jako dumny posiadacz własnego homepage’a, nieuchronnie pozostawia po sobie elektroniczne ślady.

Okoliczność tę wykorzystują twórcy wyszukiwarek adresów e-mailowych. Jednym z najskuteczniejszych przedstawicieli tej kategorii jest WhoWhere.

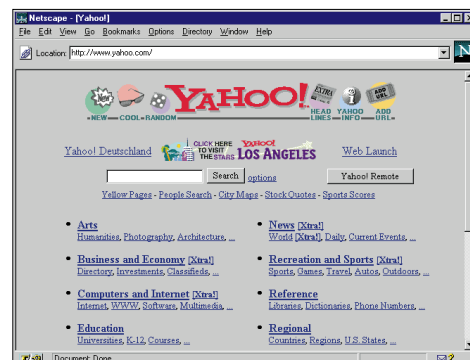
Na podstawie imienia i nazwiska WhoWhere próbuje odszukać internetowy adres danego użytkownika. E-mail można próbować odszukać nawet wtedy, gdy znane są tylko fragmenty nazwiska. Oczywiście w takim przypadku rezultat wyszukiwania jest odpowiednio mniej dokładny. Lista wyszukanych adresów może liczyć do 500 pozycji. Zaawansowane techniki wyszukiwania w przypadku tej wyszukiwarki nie są dostępne. Użytkownik może jedynie określić, czy chce otrzymywać tylko dokładne (exact) trafienia, czy także adresy w jakiegokolwiek mierze odpowiadające zadanemu pytaniu.

### Yahoo

<http://www.yahoo.com>



Yahoo kryje w sobie znacznie więcej niż mogłoby się to wydawać na pierwszy rzut oka. Niepozorny link *options* oprócz przeszukiwania Weba, umożliwia też po-



**Yahoo: orientację w cyberprzestrzeni ułatwia także przejrzysty spis kategorii tematycznych**

szukiwanie adresów e-mail i penetrowanie grup tematycznych. Do wyszukiwania można użyć operatorów OR i AND oraz określić czas utworzenia wiadomości, które mają być znalezione.

Jednym z atutów Yahoo jest możliwość przeglądania strukturalnie uporządkowanej oferty WWW. Na stronie głównej znajduje się 14 kategorii głównych – od „Arts” po „Society and Culture” – które z kolei dzielą się na dalsze podkategorie. Wiedząc do jakiej dziedziny należą poszukiwane strony można zaoszczędzić na czasie poprzez ograniczenie wyszukiwania do jednej z głównych kategorii.

*oprac. Jaromir Łański (jg)*



**Lycos: szybka wyszukiwarka udostępniająca opcje wyszukiwania, których nie oferują konkurenci**

wiązuje ten problem: wystarczy w polu do wyszukiwania wpisać „Pasadena Passadena USA” i w opcjach wyszukiwania ustawić *match 2 terms*.