



Wer suchet, der ...

Suchmaschinen im Test**Rainer Gievers/mk**

Suchmaschinen und Web-Kataloge versprechen, Ihnen im chaotischen Web den Weg zu weisen und dabei auch die sprichwörtliche Stecknadel im Heuhaufen zu finden. Dass diese Versprechen nur im Ausnahmefall auch eingehalten werden, zeigt unser Vergleichstest mit den bekanntesten deutschsprachigen Suchmaschinen

Unter der selbstgestellten Aufgabe vieler Suchmaschinen, das Web vollständig katalogisieren zu wollen, steckt eine wahre Sisyphus-Arbeit. Neben den pausenlos hinzukommenden und wegfallenden Seiten müssen auch die bereits in die Datenbank der Suchmaschine aufgenommenen Seiten regelmäßig auf Änderungen überprüft werden. Ein Suchdienst wird daher nur dann vernünftige Ergebnisse liefern, wenn die dahinterstehende Technik vernünftig konzipiert ist. Daher stellen wir dem eigentlichen Test einige Bemerkungen zu den Interna der Suchmaschinen voran.

Auf die Crawler kommt es an

Für das Aufsuchen und Indizieren der Web-Seiten werden von den Suchdiensten Crawler-Programme eingesetzt, die sich über die Links auf den Web-Seiten vorwärtsbewegen. Seiten, auf die keine externen Links ver-

weisen, fallen somit durch das Raster - Gleichermaßen gilt auch für Seiten, die mit Perl- oder PHP-generiert wurden, weil die meisten Suchmaschinen keine URLs mit CGI-Parameter aufnehmen. Deshalb werden mitunter umfangreiche Unternehmens-, Datenbanken- und Web-Shopping-Sites nicht erfasst. Wir empfehlen deshalb den Einsatz eines gut sortierten Web-Katalogs wie Web.de oder Yahoo, wenn Sie beispielsweise nach bestimmten Shopping-Sites suchen. Die Nichtaufnahme von Web-Seiten

in die Suchmaschinen ist allerdings manchmal auch vom jeweiligen Web-Site-Betreiber gewollt.

In Frame-Technik realisierte Homepages sind für einige Crawler, unter anderem von Excite und Lycos, unüberwindliche Hindernisse. Da die Hauptseite, auf der die Frames zusammengefasst werden, häufig nur einen Hinweis enthält, dass man einen Frame-unterstützenden Browser verwenden solle, geht die Homepage unauffindbar in der Suchmaschinendatenbank verloren. Falls Sie selber für Ihre Web-Seiten Frames verwenden, sollten Sie zum einen im </NOFRAMES>-Abschnitt eine Beschreibung unterbringen, worum es auf der Homepage geht, und zum anderen einen Link auf einen ungefahmten Bereich Ihrer Site setzen. Es ist in diesem Zusammenhang auch sinnvoll, auf den gefahmten Seiten Navigationselemente zu setzen, damit Besucher, die direkt von Suchdiensten auf die Seiten gelangen, sich zurechtfinden.

Einige Crawler, unter anderem von Excite und Lycos, werten keine Image-Maps aus. Es ist daher kein Wunder, dass viele Web-Sites auf Image-Maps komplett verzichten, beziehungsweise zusätzlich textbasierte Links einbauen. Nebenbei kommen die Betreiber auf diesem Wege auch Besuchern mit weniger leistungsfähigen Browsern (Handheld-Browser, textbasierte Browser wie Lynx) entgegen, die keine Image-Maps unterstützen. Um Ihre Image-Maps-orientierten Web-Seiten dennoch aufzunehmen zu lassen, können Sie

The screenshot shows a search results page from the Excite search engine. The URL in the address bar is <http://www.excite.de/search?search=the+web&submit=1&ctg=10&lang=de&cat=0>. The page title is "Excite Suchergebnisse: the web - Netscape". The search term "the web" is entered in the search field. The results include several links, one of which is for "ADVANCE ZinsGiro BANK". The page also contains instructions for using the search function, such as "Bitte beachten Sie die folgenden Tips, bevor Sie Ihre Suche wiederholen:" and "Schreibweise".

Auch wenn einige Suchmaschinen stolz vermelden, alle indizierten Web-Seiten im Volltext in ihre Datenbanken aufzunehmen, filtern viele „Stoppwörter“ heraus. Die Suche nach „the web“ bleibt deshalb in Excite erfolglos

AltaVista lieferte zu „Feigenbaum Pflege“ nur eine Web-Adresse, welche aber keine Wünsche offen ließ



natürlich auch jede der betroffenen Seiten einzeln bei den sagten Suchdiensten anmelden.

Letztendlich hängt die Qualität einer Suchmaschine sehr davon ab, welche Bestandteile der Web-Seiten in die Datenbank übernommen werden. Während sich einige auf die Auswertung der Meta Tags „Keywords“, „Description“ und den Titel beschränken, übernehmen andere, wie Fireball und Infoseek, die Seiten im Volltext. Weitere, darunter Excite, filtern häufig vorkommende Wörter, beispielsweise „the“ oder „web“ in der Annahme heraus, dass die spätere Suche danach wegen der hohen Trefferzahl keine brauchbaren Ergebnisse bringen würde.

Web-Kataloge

Die Web-Kataloge Dino-Online, Yahoo und Web.de speichern nur den Titel und eine vom Anmelder erstellte Beschreibung. Auch wenn das zunächst sehr mager klingt, liefern die Web-Kataloge wegen ihrer redaktionell gepflegten Kategoriestruktur häufig den Suchmaschinen ebenbürtige Ergebnisse und erfreuen sich folgerichtig einer großen Beliebtheit. Da die Kataloge keine Seiten selbstständig aufnehmen, ist für den Homepage-Besitzer die Anmeldung bei den wichtigsten Katalogen wie Yahoo, Web.de oder Dino-Online Pflicht.

Alle großen Kataloge leiten Suchabfragen optional an Suchmaschinen weiter, um in Fällen, in denen keine Treffer im Katalog gelandet wurde, nicht ergebnis-

nilos dazustehen. So arbeitet beispielsweise Web.de mit AltaVista zusammen.

Ranking

Wenn Sie bereits verschiedene Suchdienste mit den gleichen Begriffen „gefüttert“ haben, werden Sie sich sicher schon über die unterschiedlichen Ergebnisse gewundert haben. Für das Ranking werten die meisten Suchdienste die Häufigkeit aus, mit der die Suchbegriffe in den gespeicherten Web-Seiten vorkommen. Eine wichtige Rolle können aber auch das Verhältnis zwischen der Wörteranzahl im Verhältnis zu der Anzahl der Suchbegriffe auf der Seite, die Anzahl der Suchwörter einer Seite im Vergleich zu den anderen Seiten oder das erste Auftreten des Suchworts in der Seite, spielen. Weitere Seitenelemente, die das Ranking beeinflussen können, sind die Meta Tags, Titel, Überschriften und die URL. Für den Fall, dass Sie Optimierungen an den eigenen Web-Seiten vornehmen möchten, haben wir die Ranking-Mechanismen (soweit bekannt) in die Testtabelle übernommen. Dabei handelt es sich allerdings nur um Anhaltspunkte, weil die Suchdienstbetreiber in regelmäßigen Abständen die Ranking-Algorithmen ändern, um Manipulationen („Spamming“) zu verhindern.

Der Test

■ Accon

Accon kann nur in Bezug auf seine aufgeräumte Benutzeroberfläche punkten, während es mit der Brauchbarkeit der Sucher-

So testet die PC INTERN Suchmaschinen

Ausgewählt wurden für den Test deutsche und amerikanische Suchmaschinen, die über einen hohen Bekanntheitsgrad verfügen. Themenbezogene Suchdienste und Web-Kataloge wurden nicht berücksichtigt. Zunächst haben wir die allgemeine Bedienung unter die Lupe genommen, bei der neben der übersichtlichen Präsentation (möglichst ohne Überfrachtung mit Werbebanner) auch das Format der angezeigten Suchergebnisse eine Rolle spielte. Abgewertet wurden Suchmaschinen, deren Ausgabeformat sich nicht ändert ließ, um beispielsweise eine Kurzzeige durchzuführen sowie Suchmaschinen, die bei den Treffern nicht das Aufnahme- beziehungsweise Änderungsdatum der jeweiligen Seiten anzeigen. Gerade Letzteres wird häufig vom Nutzer benötigt, um die Aktualität der Suchergebnisse abschätzen zu können. Die Online-Hilfe sollte möglichst detailliert die Suchoptionen darlegen und Tipps zur effizienten Suche geben. In die Endabrechnung ging die Bedienung mit 30 Prozent ein. Suchmaschinen müssen so elementare Funktionen wie Phrasensuche oder zeitliche Eingrenzung unterstützen, da andernfalls keine effiziente Recherche möglich ist. Spezialoptionen, die allerdings nur selten genutzt werden, wurden zusammengefasst unter „zusätzliche Sucheingrenzung“ und „sonstige Suchoperatoren“ bewertet. In unsere Wertung haben wir nur solche Optionen aufgenommen, die in der jeweiligen Online-Hilfe dokumentiert sind. Anschließend wurden die Suchmaschinen einem 4-stufigen Test unterzogen:

1. Die Aktualität der Datenbasis wurde anhand der Suche nach „Egypt Air Absturzursache“ und „Mannesmann Vodafone“ ermittelt. Ausgewertet wurde jeweils das Datum der ersten 20 ausgeworfenen Treffer, wobei mit dem „+“-Suchoperator versucht wurde, das Ergebnis zu verbessern. Ein „Ungenügend“ bekamen alle Suchdienste, die bei einer der beiden Recherchen kein Ergebnis lieferten.

Bewertung:

>= 30 Tage: Note 6; < 25 Tage: Note 5; < 20 Tage: Note 4; < 15 Tage: Note 3; < 10 Tage: Note 2; < 5 Tage: Note 1

2. Die Fähigkeit der Suchdienste, für zusammenhängende Begriffe auf Anhieb vernünftige Ergebnisse zu liefern, wurde mit „Einstein Lebenslauf“ und „Feigenbaum Pflege“ ermittelt. Es handelt sich also um die Recherche nach jeweils zwei zeitlosen Begriffen, wie sie häufiger im Alltag vorkommt. Ausgewertet wurden die ersten zwanzig ausgeworfenen Treffer. Aus dem gefundenen Einstein-Lebenslauf mussten die wichtigsten Lebensstationen hervorgehen, während wir bei der Feigenbaumpflege zufrieden waren, wenn etwas zur Bewässerung und zum optimalen Aufstellort ausgeworfen wurde. Als Treffer wurden auch gefundene Seiten gewertet, die direkt das Gesuchte verlinkten.

3. Um Anhaltspunkte für den Umfang des Datenbestands zu erhalten, starteten wir die Anfragen „Untauglichkeitsbescheinigung“, „Monitorhelm“ und „Seitenbestandteile“. Die Wahl fiel bewusst auf diese ungewöhnlichen Begriffe, um die Ranking-Mechanismen der Suchdienste außer Kraft zu setzen, denn einige wie „Crawler.de“ bewerten auch Links, die ein Suchwort enthalten, als Treffer.

Bewertung:

0 Treffer: Note 6; < 5 Treffer: Note 5; < 10 Treffer: Note 4; < 15 Treffer: Note 3; < 20 Treffer: Note 2; ab 20 Treffer: Note 1

4. Als letzter Punkt stand die Ermittlung der „Datenbestandsfrische“ auf dem Plan. Dazu wurden die toten Links in den jeweils ersten 20 Recherche-Ergebnissen für „Seltenheitswert“, „Suchergebnisse“ und „Trauerspiel“ ermittelt. Bewertung:

< 2: Note 1; < 4: Note 2; < 6: Note 3; < 8: Note 4; < 10: Note 5; >= 10: Note 6

gebnisse eher düster aussieht. Die Online-Hilfe ist kurz und übersichtlich, was vor allem darauf zurückzuführen ist, dass nur wenige Suchoperatoren unterstützt werden. Vom angepriesenen Datenbestand, der angeblich 23,2 Millionen Web-Seiten umfasst, konnten wir während des Tests nichts merken, da weder die Aktualitätssuche noch die Recherche nach „Einstein“ beziehungsweise „Feigenbaum“ Ergebnisse brachte. Acoo erreichte daher in unserem Test den letzten Platz.

■ Aladin

Eine der hässlichsten Startseiten in unserem Testfeld besitzt Aladin, was vor allem an den vielen störenden Werbebanner liegt. Allein die Wertung der Bedienung rettet diese Suchmaschine knapp vor einem „Mangelhaft“ in der Endabrechnung.

Fazit: Die Suchmaschine Aladin ist genauso unbrauchbar wie Acoo und nicht empfehlenswert.

■ AltaVista

Beim Klassiker unter den Suchmaschinen kann zwischen einer normalen Suche und der Profisuche umgeschaltet werden. Im Gegensatz zu Fireball, das auf der gleichen Search-Engine basiert, ist es aber unerlässlich, die Online-Hilfe zu lesen, um die Vorzüge dieser Suchmaschine voll auszureißen. Ausländische Seiten lassen sich per Mausklick in eine beliebige andere Sprache übersetzen. Während Suchoptionen und Bedienung uns voll

überzeugten, kann man das von der Datenbestandsaktivität nicht behaupten, so dass sich AltaVista letztendlich mit einer guten Bewertung zufrieden geben muss.

■ Crawler.de

Nach Angaben des Betreibers zählt Crawler.de zu den ältesten deutschen Suchmaschinen. Leider wird die hakelige Bedienung an keiner Stelle erläutert, und auch die gebotenen Suchoperatoren lassen kaum eine vernünf-

Die Testergebnisse

	Acoo	Aladin	AltaVista	Crawler.de	Eule
Betreiber URL (www.)	Schöbel & Poller GbR acoon.de	Endemann!! Internet AG aladin.de	AltaVista altavista.de	Schlund und Partner AG crawler.de	Endemann!! Internet AG eule.de
Anzahl Einträge ca. Mio. deutsch / international (SearchEngineWatch)	23,2	3	16 / 275 (150)	4,8	7,3
Monatliche Seitenabrufe	1 Mio.	4 Mio.	38,6	k.A.	2
Zeitabstand Seitenanmeldung bis Index-Aufnahme	2 bis 3 Tage	ca. 3 Tage	24 h bis 1 Woche	k.A.	ca. 3 Tage
Zeitabstand zwischen zwei Crawler-Besuchen	ca. 2 Wochen	1 bis 2 Monate	abhängig von Seitenänderungsfrequenz	k.A.	1 bis 2 Monate
Seitenendungen	alles außer Binärdateien	zahlreiche Einschränkng.	k.A.	k.A.	zahlreiche Einschränkng.
Aufnahme CGI-Scripte	nein	nein	nein	nein	nein
Seitenbestandteile	Volltext	Volltext	Volltext	Volltext (max. 64 KB)	Volltext
Ranking	k.A.	k.A.	Meta Tag „Keywords“, Titel, Überschriften, Position und Häufigkeit im Text, URL, Link-Popularität, usw.	Domain, URL, Titel, Meta Tags, Pfadtiefe, Sprachqualität	Titel, Meta Tags „Keywords“, URL
Robots Exclusion Protocol	nein (in Vorbereitung)	ja	ja	ja	ja
Frames unterstützt	nein	ja	ja	ja	ja
Imagmaps unterstützt	nein	nein	ja	ja	nein
Bedienung (30%)					
Übersichtlichkeit	2	5	1	3	5
Bedienung	4	3	2	4	3
Online-Hilfe	3	2	1	5	3
Wertung	befriedigend (3)	befriedigend (3,3)	sehr gut (1,3)	ausreichend (4)	ausreichend (3,7)
Suchoptionen (30%)					
Phrasen	nein	nein	ja	ja	nein
zeitliche Eingrenzung	nein	nein	ja	nein	nein
Eingrenzung auf deutschsprachige Seiten	-	-	ja	-	-
Platzhalter (Wildcards)	nein	nein	ja	nein	nein
Zusätzliche Sucheingrenzung auf	nein	Title, URL	Domain, URL, URL-Text, Bild, Titel	nein	Title, URL, Überschriften
Sonstige Suchoperatoren	nein	nein	NEAR, Klammerung	NEAR	-
Wertung	ungenügend (6,0)	mangelhaft (5,0)	sehr gut (1,0)	ausreichend (4,0)	ausreichend (4,0)
Suchergebnisse (40%)					
1. Aktualität	keine Treffer / 6,0	keine Treffer / 6,0	ein Treffer/ 6,0	21 Tage / 5,0	22 Tage / mangelhaft
2. Suche (Aufwertung)	keine Treffer	keine Treffer	beide gefunden (+0,5)	beide gefunden (+0,5)	beide gefunden (+0,5)
3. Umfang	3 / mangelhaft	0 / ungenügend	24 / sehr gut	13 / befriedigend	3 / mangelhaft
4. Tote Links	12 / ungenügend	6 / ausreichend	4 / befriedigend	2 / gut	0 / sehr gut
Wertung	ungenügend (5,7)	mangelhaft (5,3)	befriedigend (2,8)	befriedigend (2,9)	befriedigend (3,2)
PC INTERN-Wertung	mangelhaft (5,1)	ausreichend (4,4)	gut (1,8)	ausreichend (3,6)	ausreichend (3,6)

tige Recherche zu. Durchwachsen sind auch die Suchergebnisse, die aufzeigen, dass Crawler.de zwar über eine ausreichend große Datenbasis verfügt, diese aber offenbar nicht häufig aufgefrischt wird.

■ Eule

Der Preis für das schlechteste Design geht an Eule, wo unzählige bunte, blinkende Banner den Blick des Betrachters auf sich ziehen. Wie die Recherche-Ergebnisse zeigten, kann Eule

aber gegen die Konkurrenz von Fireball & Co. bestehen, wenn auch der Datenbestand nicht besonders aktuell ist. Mängel bei Bedienung und Suchoptionen ziehen die Wertung herunter. Als einzige Suchmaschine im Test lieferte Eule keine toten Links in den Suchergebnissen.

■ Excite

Obwohl Excite mit 125 Millionen erfassten Seiten zu den größeren Suchmaschinen zählt, verfügte sie im Suchtest. Weil Excite

te in den anderen Disziplinen auf befriedigende Noten kommt, reicht es aber im Gesamtergebnis zu einem „Befriedigend“. Der Umfangstest zeigt, dass diese Suchmaschine nur vergleichsweise wenig deutsche Seiten gespeichert hat. Dies erklärt vielleicht die schlechten Suchergebnisse.

■ Fast

Fast ist zwar die größte Suchmaschine im Test, sie zeigt aber Schwächen bei den gebotenen

Suchoptionen und der Bedienung. Zusammen mit Fireball, AltaVista und Northern Light besitzt Fast offenbar die meisten deutschsprachigen Seiten im Index, nur wird dieser anscheinend nicht besonders häufig aufgefrischt, wie die Aktualitätsuche und die Anzahl der toten Links zeigten. Fast könnte für einen Anwender, der bei den anderen großen Suchmaschinen keinen Erfolg hatte, durchaus interessant sein, um doch noch einen Treffer aufzuspüren. Der Durchschnittsanwender ist dagegen



Excite	Google	Fast	Fireball	Hotbot	Infoseek	Lycos	Northern Light
Excite Inc	Google Inc.	Fast Search ASA	Gruner + Jahr	k.A.	WSI	Bertelsmann	NLTechnology LLC
excite.de	google.com	alltheweb.com	fireball.de	hotbot.com	infoseek.de	lycos.de	northernlight.com
k.A. / k.A. (125)	k.A. / k.A. (85)	k.A. / k.A. (200)	8,2	k.A. / k.A. (110)	k.A. / k.A.	k.A. / k.A. (50)	8 / k.A. (170)
k.A.	k.A.	k.A.	50 Mio.	k.A.	k.A.	k.A.	k.A.
k.A.	k.A.	k.A.	max. 24 h	k.A.	k.A.	k.A.	einige Tage
k.A.	k.A.	3 Wochen	täglich bis max. monatlich	abhängig von Seitenänderung	k.A.	abhängig von Seitenänderung	abhängig von Seitenänderung
k.A.	k.A.	Einschränkung.	Einschränkng.	k.A.	k.A.	k.A.	html, ascii, pdf
k.A.	k.A.	nein	ja	k.A.	k.A.	k.A.	nein
Volltext	Volltext	Volltext, o. Meta	Volltext	Volltext	Volltext	Volltext, o. Meta	Volltext, o. Meta
Bewertung von selbsterstellter Textzusammenfassung, Link-Popularität	Abstand der Suchbegriffe zueinander, Link-Popularität	k.A.	Seitentitel, URL, Überschriften, Häufigkeit des Suchworts im Text	Titel, Meta Tags, Anzahl der Suchwörter geteilt durch Wörteranzahl auf Seite	k.A.	k.A.	Worthäufigkeit auf Seite, Kontext, Links, Datum, usw.
ja	ja	ja	ja	ja	ja	ja	ja
nein	ja	ja	ja	nein	k.A.	nein	ja
nein	ja	nein	ja	nein	k.A.	nein	ja
4	3	1	2	2	4	3	3
3	4	4	1	3	2	3	3
1	2	4	1	1	2	2	1
befriedigend (2,7)	befriedigend (3)	befriedigend (3)	sehr gut (1,3)	gut (2)	befriedigend (2,7)	befriedigend (2,7)	gut (2,3)
ja	ja	ja	ja	ja	ja	ja	ja
nein	nein	nein	ja	ja	nein	nein	ja
ja	nein	nein	-	ja	ja	ja	ja
nein	nein	nein	ja	ja	nein	ja	ja
Einschränkungen auf Fremdsprache	nein	nein	zahlreiche Einschränkungen	zahlreiche Einschränkungen	zahlreiche Einschränkungen	zahlreiche Einschränkungen	Titel, URL, Sprache, fremdspr. Domains
Klammerung	URL-Links	nein	NEAR, Klamm.	NEAR, Klamm.	nein	NEAR, Klamme.	nein
befriedigend (2,5)	ausreichend (3,5)	mangelhaft (5,0)	gut (2,0)	sehr gut (1,0)	befriedigend (3,0)	gut (2,0)	gut (1,5)
kein Treffer / 6,0 (+0,2)	kein Treffer / 6,0 beide gefun.(+0,5)	kein Treffer / 6,0 beide gefun.(+0,5)	2 Tage / 1,0 beide gefun.(+0,5)	kein Treffer / 6,0 beide gefun.(+0,5)	5 Tage / gut beide gefun.(+0,5)	kein Treffer / 6,0 beide gefun.(+0,5)	8 Tage / gut beide gefun.(+0,5)
8 / ausreichend	2 / mangelhaft	24 / sehr gut	24 / sehr gut	3 / mangelhaft	22 / sehr gut	20 / gut	26 / sehr gut
9 / mangelhaft	4 / befriedigend	7 / ausreichend	3 / gut	2 / gut	3 / gut	13 / ungenügend	4 / befriedigend
mangelhaft (4,8)	ausreichend (4,2)	befriedigend (3,2)	sehr gut (0,8)	ausreichend (3,8)	sehr gut (1,2)	ausreichend (4,2)	gut (1,5)
ausreich. (3,8)	ausreich. (3,7)	ausreich. (3,6)	sehr gut (1,3)	befried. (2,5)	gut (2,2)	befried. (3,1)	gut (2,0)

INTERNET INTERN

Test: Suchmaschinen

Suchmaschinen an die Hand genommen

Obwohl Suchmaschinen wie eine Black Box arbeiten, haben Homepage-Betreiber erheblichen Einfluss darauf, welche ihrer Seiten dort gelistet werden. Die Notwendigkeit, Seiten von der Listung auszuschließen, kann sich beispielsweise aus der Frame-basierten Struktur der eigenen Web-Site ergeben, wenn einige Seiten nicht direkt aus einem Suchdienst heraus aufrufbar sein sollen.

In manchen Fällen ist es auch sinnvoll, Crawler bestimmter Suchmaschinen auszuschließen, wenn diese zum Beispiel Links auf CGI-Skripte verfolgen und damit die Administration der Site durcheinanderbringen. Es wäre allerdings Leichtsinn, sicherheitsrelevante oder private Seiten einfach durch das Aussperren von Suchmaschinen zu schützen, nicht nur weil einige Crawler sich über die nachfolgend beschriebenen Ausschlussmechanismen hinwegsetzen.

Robots Exclusion Protocol

Über das Robots Exclusion Protocol lassen sich einzelne Crawler von der kompletten Web-Site oder einzelnen Verzeichnissen fernhalten. Dazu muss sich im Wurzelverzeichnis der Site die Datei robots.txt befinden, weshalb sich diese Methode nicht bei Tripod, FortuneCity oder anderen Homepage-Vermarktern, die keinen Zugriff auf das Wurzelverzeichnis zulassen, anwenden lässt.

Um alle Crawler fernzuhalten, sieht der Inhalt der Datei robots.txt folgendermaßen aus:

```
User-agent: *
Disallow: /
```

Weitere Informationen zum Robots Exclusion Protocol liefert <http://info.webcrawler.com/mak/projects/robots/robots.html>. Dort wird auch eine Liste der Crawler-Namen mit Kontaktadressen der Betreiber geführt, an die Sie sich in Problemfällen wenden können.

Robots Meta Tag

Das Robots Meta Tag legt für einzelne Web-Seiten fest, ob dessen Links verfolgt werden und ob die Seite in die Suchmaschine aufgenommen werden soll. Obwohl damit eine wesentlich feinere Einteilung möglich ist, welche Seiten von den Suchdiensten aufgenommen werden sollen als mit robots.txt, greifen nur die wenigsten Crawler darauf zurück.

Ein weiterer Nachteil ist der Aufwand, den eine nachträgliche Änderung des Robots Tags auf umfangreichen Web-Sites nach sich zieht.

Zwei Beispiele für das Robots Meta Tag:

```
<meta name="robots" content="noindex, nofollow">
<meta name="robots" content="index, follow">
```

Informationen zum Robots Meta Tag sind unter <http://info.webcrawler.com/mak/projects/robots/meta-user.html> zu finden.

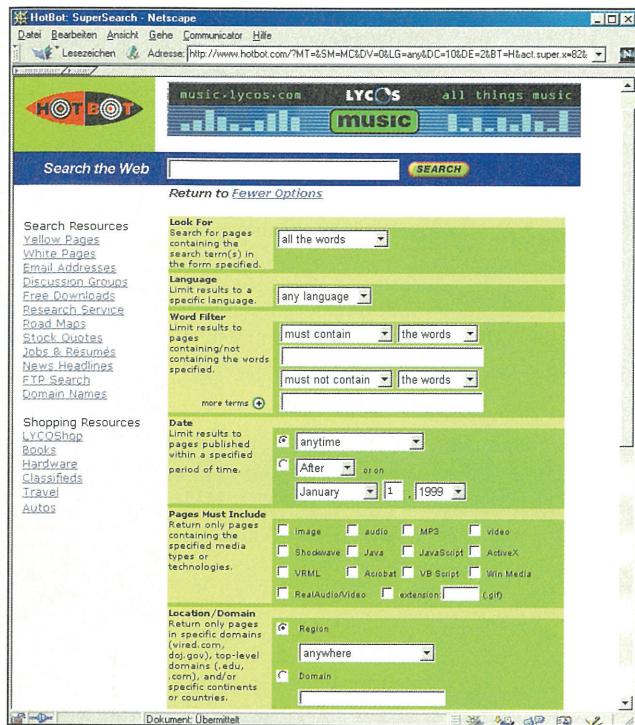
Meta-Suchmaschinen

Meta-Suchmaschinen verfügen über keine eigene Datenbasis, sondern leiten Abfragen einfach an andere Suchmaschinen oder Web-Kataloge weiter und geben deren Resultate dann zusammengefasst aus. Weil sich die Datenbasis der verschiedenen Suchmaschinen kaum überschneidet, sind Meta-Sucher besonders für Abfragen geeignet, bei denen konventionelle Suchmaschinen nur wenige oder keine Treffer liefern. Von Nachteil sind die eingeschränkten logischen Operatoren der Meta-Sucher, die sich am kleinsten gemeinsamen Nenner der genutzten Search-Engines orientieren und deshalb meist kaum über „And“ und „Or“ hinausgehen. Weitere Probleme bereitet das Ranking, welches, sofern überhaupt ausgeworfen, von den grundlegenden Suchmaschinen abhängt sowie die Beschränkung auf die ersten Ergebnisseiten der Suchmaschinen. Hilf-

reich sind Meta-Sucher aber auf jeden Fall, um Suchdienste zu einem Thema aufzuspüren, die dann direkt angesprungen werden. Wir haben für Sie die wichtigsten Meta-Sucher zusammengestellt, wobei wir nur solche berücksichtigt haben, die doppelte URLs herausfiltern. Zu beachten ist, dass deutsche Meta-Suchmaschinen meist Suchdienste des hiesigen Sprachraums auswerten, während ausländische Meta-Suchmaschinen fast ausschließlich auf amerikanische Quellen zurückgreifen.

* Logische Operatoren werden, falls von einer Suchmaschine nicht unterstützt, passend konvertiert oder herausgefiltert ; Z = Suchzeit wählbar; R = Anzahl anzuzeigender Resultate wählbar; S = Auswahl zu nutzender Suchmaschinen; + = weitere, themenspezifische Suchmöglichkeiten, beispielsweise nach News, Dateien oder den Börsenkursen.

	Name	Suchoptionen	Anzahl Suchdienste	Optionen
www.highway61.com	Highway 61	and, or	5	Z, R, gewählte Einstellungen werden gespeichert
www.mamma.com	Mamma	and, or, not (*)	8+	Z, R, S
www.metacrawler.com	MetaCrawler	and, or, phrase, regionale Einschränkung	11+	Z, R Anzeige Suchergebnisse nach Relevanz, Site, Quelle
www.profusion.com	ProFusion	and, or, not, near, Klammerung, Phrasen (*)	9+	S, R, Linküberprüfung
www.savvysearch.com	Savvy Search	and, Phrase	100+	S, Ranking der ausgewählten Suchmaschinen einstellbar,
www.apollo7.de	Apollo 7	k.A.	12	S, Z
www.multimeta.de	Multimeta	k.A.	10	S, Z
www.metager.de	MetaGer	and, or, not	21	S, R, Z, Relevanzanalyse, Linküberprüfung, Sortierung
www.metaspinner.de	Metaspinner	and, or, not, Phrasen *	16+	S
www.netz.de	netz.de	and, or, Phrasen	21+	S, Linküberprüfung
www.suchen.com	suchen.com	and, or	40	S, Z, R
www.topXplorer.de	topXplorer	and, or, not, Phrasen	5+	Relevanzanalyse



Für komplexere Suchabfragen stellen viele Suchmaschinen eine Extra-Maske bereit, so dass sich der Nutzer das Erlernen komplexer Suchoperatoren erspart

bei Fireball, AltaVista oder Northern Light besser aufgehoben.

■ Fireball

Fireball besitzt ein eigenwilliges, aber dennoch übersichtliches Design und kann in der Bedienung weitgehend überzeugen. Gleiches gilt auch für den umfangreichen Vorrat an logischen Operatoren, die in der verständlichen Online-Hilfe mit Beispielen erläutert werden. Wer sich nicht mit kryptischen Suchoperatoren auseinandersetzen will, kann auch die Detailsuche nutzen. Fireball beschränkt sich zwar standardmäßig auf deutsche Web-Seiten, kann aber auch auf den Altavista-Datenbestand zurückgreifen. Den Suchtest durchlief Fireball mit Bravour und setzte sich daher auch an die Spitze des Testfelds.

■ Google

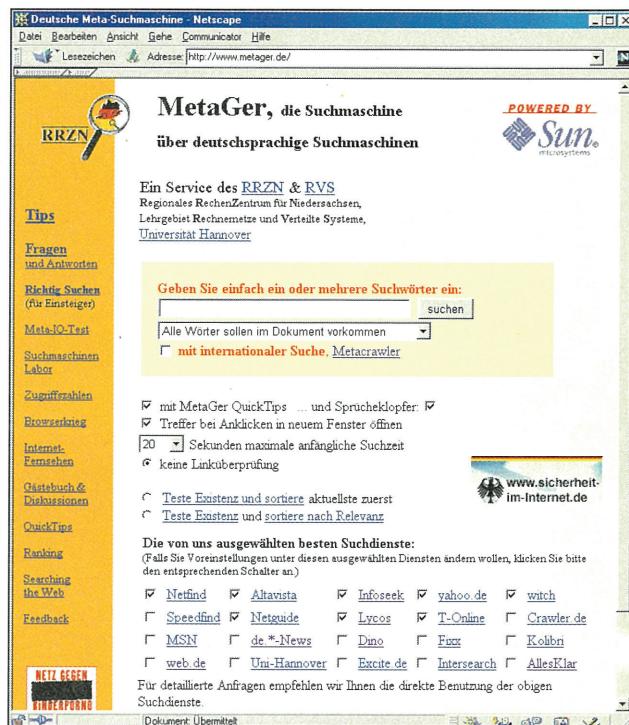
Auch dieser Suchdienst kommt nur auf eine ausreichende Bewertung, wozu die wenigen Suchoptionen und die schlechten Recherche-Ergebnisse bei-

tragen. Einmalig im Testfeld ist die Option, Web-Seiten, so wie sie in der Google-Datenbank gespeichert sind, anzeigen zu lassen, um beispielsweise doch noch zu einem Ergebniss zu gelangen, auch wenn die Originalseite nicht mehr vorhanden ist.

■ Hotbot

Der Datenbestand von Hotbot wird von Inktomi geliefert, einem Unternehmen, das keine eigene Suchmaschine betreibt, sondern seine Dienste ausschließlich an Suchmaschinenbetreiber lizenziert. Die grellbunte Hotbot-Oberfläche dürfte nicht jedermann sache sein, sie ist aber trotzdem übersichtlich und benutzerfreundlich. Für die Spezialsuche ist der Nutzer nicht auf die Eingabe von kryptischen Suchoperatoren angewiesen, sondern kann auf ein Formular umschalten, über das alle wichtigen Operatoren zur Sucheingrenzung zur Verfügung stehen.

Wegen Mängeln in den Suchergebnissen bekam Hotbot in der Endabrechnung dennoch nur einen Platz im Mittelfeld.



Meta-Suchmaschinen wie MetaGer machen sich die kaum überlappenden Datenbestände der konventionellen Suchdienste zu Nutze und liefern daher häufig auch zu schwierigen Abfragen noch Ergebnisse

■ Infoseek

Infoseek kann zwar nicht durch umfangreiche Suchoptionen oder eine tolle Bedienung auftrumpfen, überzeugte uns aber durch ausgezeichnete Recherche-Ergebnisse. Wer häufiger Suchdienste für komplexe Recherchen nutzt, sollte Infoseek auf jeden Fall einmal ausprobieren. In der Gesamtwertung belegt Infoseek einen der oberen Plätze im Testfeld.

■ Lycos

Lycos hinterließ einen durchwachsenen Eindruck, weil die logischen Operatoren eine gute und die Bedienung eine befriedigende Note erhielten, die Suchabfrage aber Mängel offenbarte, die nur eine ausreichende Note ergaben. In der Endabrechnung kam Lycos auf ein „Befriedigend“.

■ Northern Light

Northern Light kam in allen Kategorien auf eine gute Bewertung. An die 26 Treffer im Um-

fangstest konnte kein Konkurrent heranreichen, weswegen Northern Light sehr interessant für schwierige Recherchen ist. Northern Light kommt so auf den dritten Platz unseres Tests.

Fazit

Unser Test hat gezeigt, dass viele Suchmaschinen kaum ihren Namen verdienten. Wir waren jedenfalls überrascht über die Schwierigkeiten, die einige bei der Recherche nach einfachen Begriffen wie „Feigenbaum Pflege“ und „Einstein Lebenslauf“ hatten. Testsieger wurden Fireball, AltaVista und Northern Light, aber auch Infoseek hinterließ einen brauchbaren Eindruck.

Anzeige

www.computerjobs.de

Ihr
@
Stellenmarkt