

# 3 Wissenswertes zu Suchmaschinen

Suchmaschinen sind Datenbanken. Es gibt mehrere Typen von Suchmaschinen: Roboter, auch Crawler oder Spider genannt, und Verzeichnisdienste. Außerdem gibt es Meta-Suchmaschinen und Spezialsuchdienste.

## 3.1 Suchmaschinen-Typen

Im folgenden werden die verschiedenen Typen von Suchmaschinen voneinander unterschieden. Tatsächlich sind viele Suchhelfer Mischtypen oder Kombinationen.

### 3.1.1 Roboter

Suchmaschinen vom Typ Roboter sind auf extreme Hochleistung getrimmte Systeme, die laufend das World Wide Web und andere Elemente des Internet analysieren. Google beispielsweise setzt nach eigenen Angaben<sup>1</sup> einen Cluster aus über zehntausend PCs ein, die auf vier Datenzentren verteilt sind.

Roboter bestehen im Grunde aus drei Elementen:

- ▶ dem eigentlichen Crawler, der ähnlich wie ein Browser Seiteninhalte aufruft und diese dann ausliest, er findet auch Hyperlinks und kann diesen folgen, um weitere Seiten zu entdecken,
- ▶ der Datenbank, die Wortlisten führt und zu den Wortlisten Fundstellen (URLs) kennt, man bezeichnet diese Datenbank als Index,
- ▶ der Web-Oberfläche, über die der Benutzer mit der Suchmaschine interagiert.

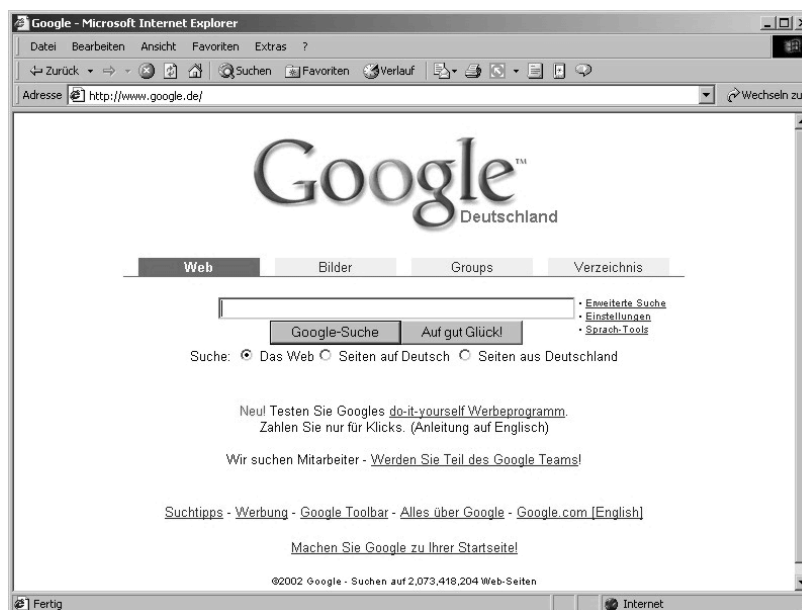
---

<sup>1</sup> Die Angaben sind einem Interview mit Monika Henzinger, der Leiterin der Forschungsabteilung von Google in Mountain View, Kalifornien, in der Zeitschrift »Spektrum der Wissenschaft« vom Juli 2002 entnommen.

### 3 Wissenswertes zu Suchmaschinen

Nach diesem Prinzip arbeiten unter anderem Google, Altavista und AOL. Crawler übergeben nicht nur die gefundenen Wörter an die Datenbank, sondern finden und lesen auch Hyperlinks und verfolgen diese weiter. Auf diesem Weg können sie zuerst die gesamte Website und von hier aus auch weitere Sites erfassen, zumindest soweit diese durch Hyperlinks gefunden werden können.

Abbildung 3.1:  
Eine Suchmaschine des Typs  
Roboter: Google



Wenn man sich allerdings überlegt, dass jede Seite auf mehrere andere Seiten verweisen könnte, die wiederum auf viele weitere Seiten verweisen und so weiter, wird einem schnell klar, dass es für solche Crawl-Verfahren Grenzen geben muss, ansonsten würde die Liste der vom Crawler noch zu besuchender Seiten ja deutlich schneller anwachsen, als er sie abarbeiten kann.

## Suchmaschinen-Typen

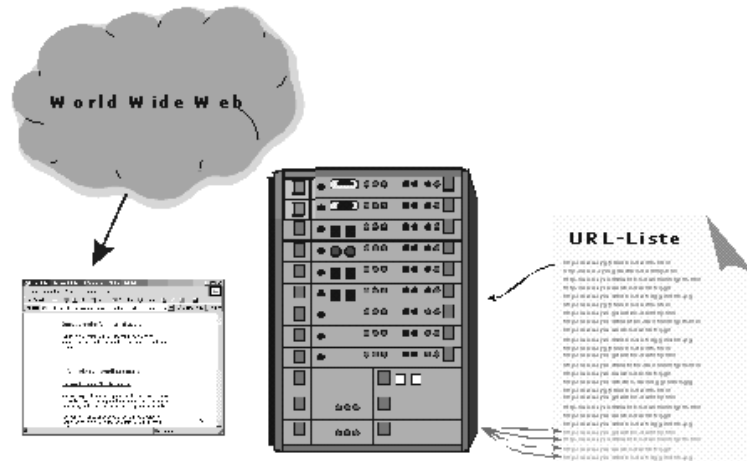


Abbildung 3.2:  
Die Suchmaschine arbeitet eine Liste mit URLs ab, fügt aber selbst laufend neue URLs hinzu; die Liste wächst unten schneller an, als sie oben abgearbeitet wird

Auch der leistungssteigernde Einsatz von schnelleren Rechnern und so genannten Clustern, mehreren parallel und voneinander weitestgehend unabhängig arbeitenden Rechnern, die ihre Ergebnisse in eine gemeinsame Datenbank übergeben, ist keine Lösung zumindest für dieses Problem: Mit der erhöhten Leistung wächst auch die Liste der noch zu besuchenden Seiten schneller. Der Einsatz solcher Cluster lässt sich bei einigen Suchmaschinen übrigens tatsächlich beobachten, speziell wenn man Protokolldaten des Webservers auswertet, worüber wir noch sprechen werden.

Die Lösung für Roboter ist es, die Liste der abzuarbeitenden URLs immer wieder zu verkürzen, also Einträge rauszulöschen. Wenn wir über den Umgang mit Suchmaschinen und die Probleme, die sich dabei ergeben können, sprechen, werden wir auf diesen Punkt zurückkommen müssen.

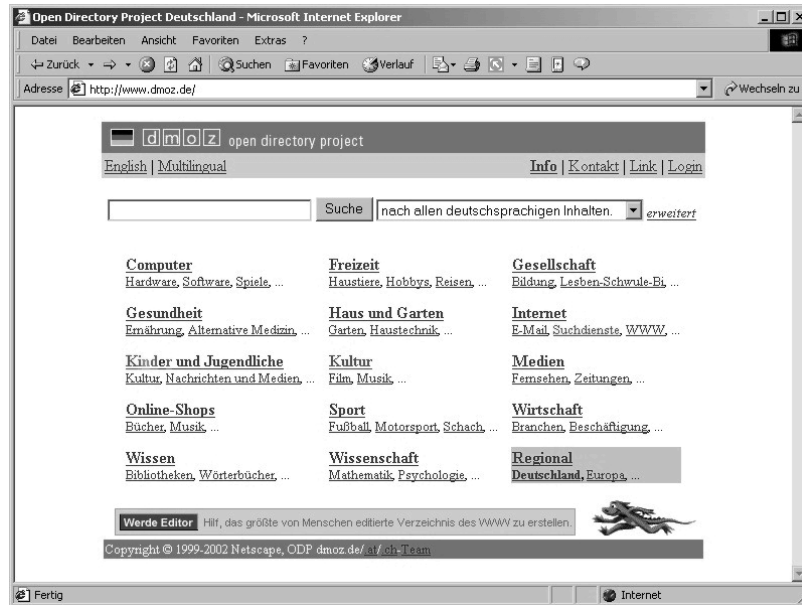
### 3.1.2 Verzeichnisdienste

Verzeichnisdienste funktionieren im Grunde ähnlich, statt eines Crawlers prüft allerdings zuerst ein Mensch die Inhalte, bewertet sie und stellt sie dann in die Datenbank. Es gibt auch einige Verzeichnisdienste, bei denen die Aufnahme eines neuen URL automatisiert erfolgt.

Bekannte Verzeichnisdienste sind Yahoo, Dino-Online, Web.de oder – über das Open Directory Project – Netscape.

### 3 Wissenswertes zu Suchmaschinen

Abbildung 3.3:  
Beispiel für einen  
Verzeichnisdienst: Das von  
Menschen redaktionell geführte  
Open Directory  
Project



Die Anzahl von Fundstellen, über die ein Roboter zu einem beliebigen Begriff verfügt, ist meist um ein Vielfaches höher, als bei einem Verzeichnisdienst, dieser verspricht dafür aufgrund der Bewertung durch Menschen eine höhere Relevanz der Ergebnisse und Ausschluss von unerwünschten Ergebnissen, z. B. von pornografischen Seiten.

Verzeichnisdienste fassen Angebote außerdem in Kategorien zusammen, die eine alternative Suchmöglichkeit zur Stichwortsuche ermöglichen. Diese Kategorien machen auch den Unterschied bei der Anmeldung aus: Während Roboter meist mit der Angabe des URL und einiger weniger zusätzlicher Infos zufrieden sind, wollen Verzeichnisdienste wissen, in welcher Kategorie die neue Website aufgenommen werden soll.

Die Kategorien der Verzeichnisdienste sind hierarchisch aufgebaut. Einen kleinen Ausschnitt der Verzeichnishierarchie von Yahoo zeigt die folgende Grafik:

### Suchmaschinen-Typen

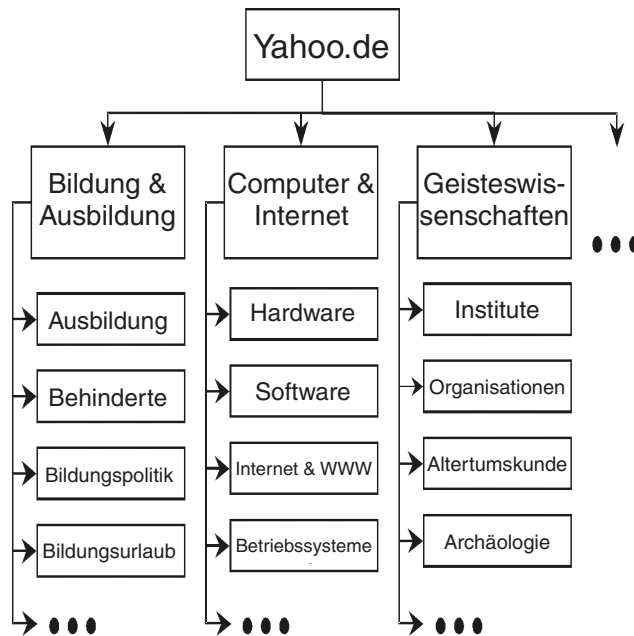


Abbildung 3.4:  
Verzeichnishierarchie bei Yahoo  
(Ausschnitt)

Sobald der Redakteur eines Verzeichnisdienstes darüber entschieden hat, dass eine Website aufgenommen wird und in welcher Kategorie, wird zusätzlich ein Roboter eingesetzt, bei Open Directory Project heißt dieser beispielsweise Robozilla. Ausgehend von der angemeldeten Startseite werden über Hyperlinks weitere Seiten gefunden und in den Index aufgenommen. Der Verzeichnis-Roboter besucht anschließend regelmäßig die Website, prüft sie auf Änderungen und informiert den Redakteur, falls sie nicht mehr verfügbar ist. Ähnlich wie bei den eigentlichen Robotern kann eine solche Situation auch beim Verzeichnisdienst dazu führen, dass die einmal angemeldete Site wieder entfernt wird.

Wenn der Benutzer den Verzeichnisdienst über den hierarchischen Katalog durchsucht, werden ihm deshalb nur die Startseite der jeweiligen Websites als Ergebnisse vorgeschlagen, von denen aus sich der Benutzer dann weiter durchklicken kann.

Webverzeichnisse bieten allerdings meist auch ein Suchfeld an, in das wie bei Robotern Suchbegriffe eingegeben werden können. Hier hat der Benutzer nun Zugriff auf den vollen Bestand des Indexes. In der Regel kann der Benutzer hier auch Ergebnisse aus Robotern finden, was insbesondere sinnvoll ist, wenn die eigenen Ergebnisse des Verzeichnisdienstes zu wenig ergiebig sind.

Webverzeichnisse werden meist von Menschen gepflegt und versprechen so, relevante Inhalte zuerst anzubieten, denen die große Menge der vom Roboter gefundenen Seiten folgt.



Verzeichnisdienste und Robots arbeiten oft eng zusammen. Während der Robot Google die Möglichkeit anbietet, alternativ auf das Verzeichnis des Open Directory Project zuzugreifen, bietet der Verzeichnisdienst von Yahoo.de an, die Daten des Robots Google mit anzuzeigen.

Eine Besonderheit ist der Verzeichnisdienst des Open Directory Project. Hier hat der Benutzer selbst die Möglichkeit, Redakteur zu werden, wobei grundsätzlich nicht einmal untersagt ist, eigene Seiten anzumelden und hoch zu bewerten. Gewarnt wird nur davor, dass es zum Ausschluss des Redakteurs führen kann, wenn eine missbräuchliche Nutzung erkennbar ist, beispielsweise Ausschluss eines Wettbewerbers.

Wenn man dann noch bedenkt, dass Google die Bewertungen des Open Directory Project in sein eigenes Ranking einfließen lässt...

### 3.1.3 Meta-Suchmaschinen

Ein dritter Typus sind Meta-Suchmaschinen, wie zum Beispiel Meta-Ger oder *www.suchen.com*. Diese besitzen keine eigenen Daten, die sie durchsuchen, sondern leiten die Anfrage im Hintergrund an mehrere andere Suchmaschinen weiter. Häufig kann der Benutzer auswählen, von welchen Suchmaschinen er die Ergebnisse sehen möchte. Da die angefragten Suchmaschinen teilweise unterschiedliche Datenbestände haben, erhält der Benutzer umfangreichere Suchergebnisse, was nicht unbedingt immer wünschenswert sein muss. Außerdem dauert es oft recht lange, bis alle Ergebnisse zusammengetragen und in einer Ergebnisseite verarbeitet sind, insbesondere, wenn man die Suchdauer mit der von Google und anderen Sprintern vergleicht.

## Suchmaschinen-Typen



Abbildung 3.5:  
Eine Meta-Suchmaschine mit einer Auswahl an Suchmaschinen, an die eine Anfrage weitergeleitet werden kann.

Ein weiterer Ansatz ist, Meta-Suchmaschinen themenorientiert einzurichten. Diese durchsuchen dann andere Suchmaschinen, filtern aber nach den für sie relevanten Informationen aus. Das könnten Gebrauchtwagen sein, Software und viele andere. Sie verbinden also die Vorteile der Meta-Suchmaschinen mit denen der im folgenden beschriebenen Spezialsuchmaschinen.

Meta-Suchmaschinen können für viele der hier beschriebenen Aufgabenstellungen recht nützlich sein, da sie einen vergleichsweise schnellen Überblick erlauben, in welchen Suchmaschinen eine Website bekannt ist, und welche Platzierung sie unter den Suchergebnissen hat. Man spart es sich damit, die Suchmaschinen mühselig einzeln abzuarbeiten.

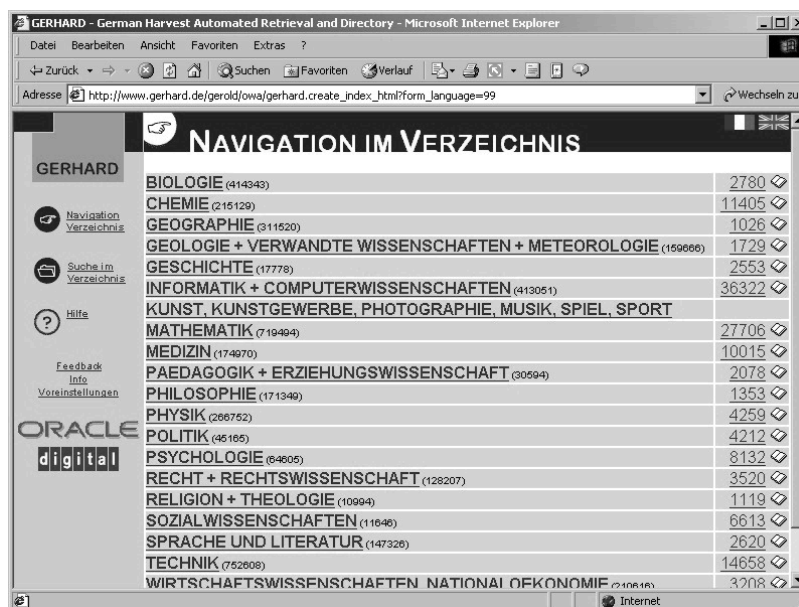
Einige Meta-Suchmaschinen bieten als Geschäftsmodell an, für Kunden die Anmeldung bei den eigentlichen Suchmaschinen durchzuführen – natürlich gegen einen kleinen Obolus.

Wenn man an solch einer Dienstleistung interessiert ist, sind die Meta-Suchmaschinen sicher keine schlechte Adresse.

### 3.1.4 Spezial-Suchmaschinen

Weiter zu berücksichtigen sind Suchseiten, die sich speziell an bestimmte Branchen richten. Auch sind Angebote, die sich auf spezielle Themen konzentrieren, hierunter zu fassen. Diese sind typischerweise redaktionell geführt und wirken gelegentlich nicht sehr professionell. Die dazu notwendige Fleißarbeit sollte aber nicht unterschätzt werden, zumal die Autoren diese Seiten selten hauptberuflich pflegen.

Abbildung 3.6:  
Eine Spezialsuch-  
maschine für  
wissenschaftliche  
Seiten: Gerhard



Diese Verzeichnisse lassen sich eventuell nutzen: Kontaktieren Sie den Anbieter und versuchen Sie, darin mit einem Hyperlink auf Ihre Website platziert zu werden.



### Suchmaschinen-Typen

Wenn Sie damit erfolgreich waren, können Sie die Seiten, auf denen der Link zu Ihnen zu finden ist, selbst bei Robotern anmelden, und wie wir sehen werden, kann dies Ihre Platzierung in den eigentlich interessierenden Suchmaschinen verbessern helfen.

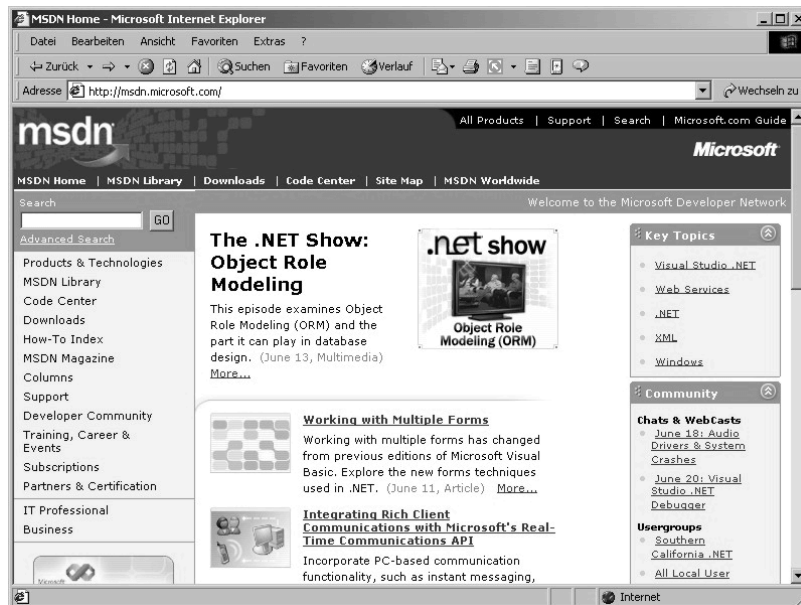
Für gewisse Branchen sind solche Spezial-Suchmaschinen bereits zu einem wertvollen und unverzichtbaren Service geworden. So gibt es für den Pressebereich Suchmaschinen, anhand derer man veröffentlichte Artikel und Nachrichten finden kann. Oder auch solche, die eine Suche speziell nach Bildern und Grafiken oder mp3-Dateien ermöglichen. Adressen finden Sie im Anhang. Oder wie wäre es mit einer kirchlichen Suchmöglichkeit?



Abbildung 3.7:  
Kirchliche Spezialsuchmaschine

Es gibt auch Suchmaschinen, die sich auf das Web-Angebot eines Anbieters konzentrieren. Microsoft beispielsweise hat mehrere solcher Suchmaschinen im Einsatz. Das folgende Beispiel durchsucht MSDN, das Microsoft Developer Network, mit einer sehr großen Anzahl von Artikeln, die insbesondere für Programmierer von Interesse sind.

Abbildung 3.8:  
Suchmaschine für  
technische Artikel  
bei Microsoft



Eine breite Übersicht über Suchmaschinenangebote finden Sie übrigens unter anderem unter [www.suchmaschinen.de](http://www.suchmaschinen.de). Roboter und Verzeichnisdienste sind zwar nicht voneinander getrennt, dafür ist die Sammlung aber recht umfangreich, auch einige Spezial-Suchmaschinen finden sich dort.

## 3.2 Die Bedeutung von Suchmaschinen für Anbieter von Web-Inhalten und eventuelle Probleme

Wahrscheinlich ist Teil Ihrer eigenen Erfahrung mit dem Internet, dass Suchmaschinen am Beginn jeder Recherche stehen. Es ist immer wieder erstaunlich, wie schnell man über sie Informationen zu Themen, Firmen, Personen oder Produkten findet, und seien sie noch so exotisch und selten.

Sie wissen, beispielsweise, dass ein wichtiger Lieferant im Internet vertreten ist und dort Datenblätter zu seinen Produkten eingestellt hat. Sie kennen zwar den URL der Firma nicht, aber das ist kein Problem. Sie rufen *Google* oder *Altavista* auf oder eine andere Suchmaschine, geben den Namen der Firma ein, und erhalten eine Liste von Dokumenten, die mit dem Unternehmen zu tun haben. Wahrscheinlich ist auch schon mindestens ein Dokument aus der Website der Firma dabei. In seltenen Fällen mag es notwendig werden, zu »blättern«, sich also weitere Seiten mit Ergebnissen anzeigen zu lassen, aber wenn die gesuchte Firma tatsächlich eine eigene Website hat, werden Sie sie mit Sicherheit finden.

Eigentlich interessiert Sie doch aber nicht der Lieferant, den kennen Sie ja bereits. Sie suchen ein Datenblatt zu einem Produkt, das Sie interessiert. Die Wahrscheinlichkeit ist groß, dass Sie das Datenblatt finden, wenn Sie gleich die Produktbezeichnung in eine Suchseite eingeben.

Vermutlich werden Sie die gewünschte Information sogar schneller finden, als wenn Sie zuerst auf die Website der Firma gehen und sich von dort aus über mehrere Hyperlinks bis zum gewünschten Inhalt durchhangeln. Probieren Sie es doch mal aus: Schauen Sie nach, ob es für Ihren Drucker beim Hersteller einen aktuelleren Treiber gibt als den, den Sie gerade einsetzen. Machen Sie das auf zwei Wegen (und stoppen Sie jedes Mal die Zeit)<sup>2</sup>:

- ▶ Rufen Sie den URL des Herstellers auf (z.B. *www.hewlett-packard.com*) und suchen Sie von der Startseite aus die Seite, von der aus Sie den Druckertreiber herunterladen können.
- ▶ Gehen Sie zu einer der großen Suchmaschinen (z.B. *www.google.de*), geben Sie die Bezeichnung des Druckers ein, dann das Wort »driver« und dann das Wort »download«.

Spätestens jetzt sollte klar sein, warum Suchmaschinen so wichtig sind und ihre Popularität eher noch zunehmen wird. Gut genutzte Suchmaschinen sind meist der schnellste Weg zur gesuchten Information.

---

<sup>2</sup> Natürlich spricht auch nichts dagegen, stattdessen die de-Toplevel-Domänen aufzurufen und nach einem deutschen »Treiber« zu suchen.

Es gibt heute sogar professionelle Trainings für Internet-Recherche. Einkäufer, Journalisten und Entscheider aus den unterschiedlichsten Bereichen lernen hier, wie man das World Wide Web und hier insbesondere die Suchmaschinen nutzt, um Informationen zu finden. Und auch Literatur gibt es hierzu.

Sie, liebe Leserin, lieber Leser, bieten doch auch etwas an. Sie haben Inhalte im Internet, von denen Sie möchten, dass Benutzer sie finden, wenn sie sie brauchen. Sie möchten gefunden werden, und zwar von den Personen, die Suchmaschinen als Recherchemittel einsetzen.

Auf den ersten Blick ist das einfach, tatsächlich stellen sich uns einige ganz vertrackte Probleme:

► Problem 1: Suchbegriffe und Schreibweisen berücksichtigen

Überlegen Sie einmal, nach welchen Stichwörtern ein Benutzer suchen könnte, von dem Sie wünschen, dass er Ihre Website findet. Vermutlich wird es dafür eine recht große Anzahl von Möglichkeiten geben, zum einen natürlich unterschiedliche Begriffe und Wörter, aber auch nur verschiedene Schreibweisen, die sich teilweise nur geringfügig unterscheiden.

Probieren Sie diese Möglichkeiten einmal mit einer Suchmaschine aus. Sie werden feststellen, dass die Ergebnisse jedes Mal recht unterschiedlich sein können. Hier ein Beispiel aus der Welt eines der Autoren:

- Projektmanagementtraining
- Projekt-Management-Training
- Projektmanagement-Training
- Projektmanagement-Seminar
- Projekt-Management-Seminare
- Projektleiter-Seminare
- Seminare für Projektleiter
- Project Management Training
- project methodology seminar
- Kurse in Project Management

Wo es verschiedene Begriffe und Schreibweisen gibt, sollten Sie damit rechnen, dass auch alle verwendet werden.

Übrigens: Beobachtungen haben ergeben, dass die meisten Benutzer von Suchmaschinen normalerweise nicht mehr als zwei Begriffe in das Suchfeld eingeben. Auch aus einem Begriff wie »Seminare für Projektleiter« wird oft nur »Seminare Projektleiter«, weil in den meisten Suchmaschinen häufig verwendete Wörter wie »für«, »und« oder »der«/»die«/»das« in der Suche weggelassen werden.

Eine Hilfe beim Finden des richtigen Suchbegriffs sind Listen von häufigen Benutzer-Eingaben.

Fireball bietet unter [www.fireball.de/qstat.html](http://www.fireball.de/qstat.html) eine solche Liste als Stichwort-Statistik, die hochaktuell ist. Sie zeigt die angefragten Begriffe des Vortages und wird laufend aktualisiert. Die Liste ist allerdings zensiert, mit Sex und Erotik verbundene Begriffe sind entfernt worden. Vermutlich wird Sie das bei Ihren Bemühungen eher unterstützen.

Eine ähnliche, jedoch unzensierte Liste findet sich bei Suchmaschinentricks.de unter [www.suchmaschinentricks.de/tools/keywords.php3](http://www.suchmaschinentricks.de/tools/keywords.php3). Die hier zu findende Liste bezieht sich nach Aussage ihres Autors auf Beobachtungen, die über einen längeren Zeitraum gemacht worden sind.

Aber Vorsicht! Bevor Sie jetzt Begriffe wie *SMS* oder *Download kostenlos* auf Ihrer Titelseite platzieren: Erreichen Sie damit wirklich die Benutzer, an denen Sie interessiert sind?

Die Website [www.suchmaschinentricks.de/tools/kw\\_lookup.php3](http://www.suchmaschinentricks.de/tools/kw_lookup.php3) bietet übrigens einen noch interessanteren Dienst an: Die Suche nach den häufigsten Kombinationen, in denen ein Wort von Benutzern gesucht worden ist. Wenn Ihre Seite zu der gesuchten Kombination eine besonders hohe Relevanz hat, wird sie bei den Suchergebnissen top platziert sein. Der Begriff Relevanz wird übrigens in Kapitel 4 ausführlich beschrieben.

Wir werden in diesem Buch auch noch darüber sprechen, wie wir feststellen können, welche Schreibweise ein Besucher verwendet hat, wenn er oder sie die Website in der Suchmaschine gefunden hat, denn eines ist klar: Gehen Sie nicht davon aus, dass es die von Ihnen im allgemeinen verwendete oder erwartete ist.



► Problem 2: Die richtigen Suchmaschinen finden

Noch etwas weiteres müssen wir in Erfahrung bringen: Welche Suchmaschinen werden von unseren potenziellen Neukunden bevorzugt? Und wie werden diese eingesetzt? Die meisten haben eine Suchmaschine, die sie am liebsten verwenden, vielleicht, weil die Ergebnisse so genau sind, oder so umfangreich, oder weil der Benutzer die vielen kleinen Tricks der Experten-suche beherrscht, die bei dieser Suchmaschine funktionieren und bei einer anderen nicht beziehungsweise mit anderen Eingaben.

Manch einer bevorzugt vielleicht Google, weil die Oberfläche so schlicht ist und sich auf das Wesentliche konzentriert, und weil der Benutzer sich nicht über lästige Pop-Up-Fenster ärgern muss. Auch verfügt Google, zumindest nach eigener Aussage, über die größte Suchwort-Datenbank (über 2 Milliarden Seiten). Andere suchen über Yahoo, Lycos oder Netscape, wo es redaktionell geführte Verzeichnisse gibt (siehe oben) und wo man nebenher weitere Informationen sammeln und Nachrichten lesen kann. Wieder andere bevorzugen Meta-Suchmaschinen wie Meta-Ger, die Such-Anfragen nicht selbst bearbeiten, sondern gleich an mehrere andere Suchmaschinen durchreichen und eine entsprechend höhere Menge an Ergebnissen haben, die aber auch im Umgang eher anspruchsvoller sind. Sie werden gerne von Studenten verwendet und von Personen, die ihr Einkommen im Bereich Forschung und Lehre erzielen.

Manche verwenden gerne national ausgerichtete Suchmaschinen, andere eher die Global Players und so weiter. Es mag noch viele andere Gründe geben, warum sich Benutzer für eine Suchmaschine entscheiden, wir können nicht alle kennen. Und wir wissen, dass sich Bevorzugungen auch wieder ändern können.

Welche Suchmaschinen werden gerne genutzt? Für Deutschland gibt es bei [www.etracker.de](http://www.etracker.de) unter dem Punkt »globale Statistik Suchmaschinen« eine Statistik für eine Reihe beobachteter Webseiten, von welchen Suchmaschinen aus Besucher zu ihnen gefunden haben. Anfang 2002 sah die Reihenfolge bei den wichtigsten Suchmaschinen wie folgt aus:

### Bedeutung und Probleme von Suchmaschinen

Rang	Suchmaschine	Anteilig
1	Google	63.7%
2	Yahoo	7.5%
3	Lycos.de	6.1%
4	T-Online	4.4%
5	Web.de	3.7%
6	MSN Germany	3.3%
7	MetaGer	2.4%
8	Altavista	1.9%
9	Fireball	1.8%
10	Freenet	1.1%
11	MSN	0.8%
12	AllesKlar	0.5%
13	Dino Online	0.4%
14	NetCraft	0.4%
15	Crawler.de	0.3%

Aber Vorsicht: Dies muss für Ihre spezifische Website und die von Ihnen gewünschte Zielgruppe nicht identisch sein. Dennoch ist der von Google erzielte Anteil von über 60 Prozent der Weiterleitungen an Websites aus Suchergebnissen heraus beeindruckend, und die Tendenz ist weiter steigend.

► **Problem 3: Die richtigen Besucher ansprechen**

Wollen Sie wirklich jeden Besucher? Stellen wir uns den Marktplatz Internet als Krämer- oder Wochenmarkt vor. Wollen Sie von jedem, der vorbeigeht, dass er/sie stehen bleibt? Auch bei Desinteresse? Irgendwann ist der Platz vor Ihrem Stand gefüllt mit Betrachtern, die eigentlichen potenziellen Kunden finden keinen Platz mehr und gehen an Ihnen vorbei.

Ähnliches kann Ihnen auch im Internet passieren. Für eine große Besucherzahl müssen Sie gerüstet sein, der Webserver muss ausreichend Bandbreite (Übertragungsleistung der Anschlussleitungen zum Internet), Festplattendurchsatz und Rechenleistung haben, um dies verarbeiten zu können. Wenn der Server in einem dieser Faktoren überlastet wird, kann dies zu unangenehmen Erlebnissen für Ihre Besucher führen: Das

Antwortverhalten des Servers wird immer träger, im schlimmsten Fall kommt er zum Stillstand. Oder der Benutzer erhält eine Fehlermeldung. Alles dies kann negativ auf Ihr Unternehmen zurückfallen.

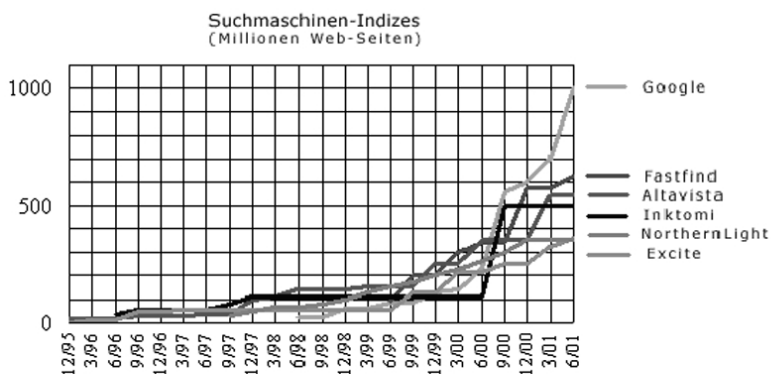
Ein guter Administrator wird die Last auf dem Server überwachen und rechtzeitig nachrüsten, bevor Engpässe bestehen, aber gute Administratoren sind eher knapp und eine Nachrüstung ist unter Umständen auch nicht billig.

Ungewollte Besucher können also auf Kosten der interessanten die Leistung reduzieren. Unabhängig davon, ob Sie eigene Server einsetzen oder einen Dienstleister nutzen: Leistungsfähigkeit im Internet gibt es nicht zum Nulltarif. Dies gilt insbesondere, wenn Sie dynamisch erzeugte Seiten anbieten, da es für den Rechner nicht reicht, diese von der Platte abzurufen und an den Benutzer zu senden. Sie müssen oft umfangreiche Bearbeitungsschritte erfahren, die den Prozessor des Servers belasten und die Antwortzeiten erhöhen.

### 3.3 Indexgrößen wichtiger Suchmaschinen

Die folgende Grafik zeigt einige große, internationale Suchmaschinen, und wie sich deren Indexgrößen, die Listen der erfassten Wörter, im Zeitraum von über 5 Jahren entwickelt haben:

Abbildung 3.9:  
Suchmaschinen-  
Indizes, Entwick-  
lung 1995 bis  
2001





### Indexgrößen wichtiger Suchmaschinen

Die Zahlen, auf denen diese Grafik basiert, stammen von den Betreibern selbst, sind also mit Vorsicht zu genießen.

Man kann deutlich erkennen, wie bei den meisten Anbietern die Indexgrößen laufend steigen. Für viele war das Jahr 2000 der Zeitraum des schnellsten Wachstums. Waagerechte Linien im Diagramm müssen nicht bedeuten, dass der entsprechende Index stagniert, sie können auch nur daher rühren, dass vom Suchmaschinenbetreiber keine neuen Angaben zu finden waren. Eine andere Interpretation ist die, dass die Erweiterung eines Index eine recht teure und technisch aufwändige Aktion zu sein scheint, die von den Suchmaschinen nur im Rahmen von gelegentlichen größeren Projekten durchgeführt werden kann. Immerhin soll ja auch während der Umstellung die Suchdienstleistung weiter ungestört angeboten werden.

Google startete erst im Juni 1998, hat aber – wenn die vom Unternehmen mitgeteilten Zahlen stimmen – mit 1 Milliarde Seiten im Sommer 2001 und einer Erweiterung auf 2 Milliarden Seiten im Spätherbst des gleichen Jahres den bei weitem umfangreichsten Index.

Eine weitere Übersicht stammt von *www.at-web.de*. Hier finden sich auch Angaben speziell zu deutschsprachigen Suchmaschinen sowie zu Verzeichnisdiensten. Die Angabe »Datum« bezieht sich auf den Zeitpunkt der Veröffentlichung.<sup>3</sup>

Suchdienst	Index	Datum	Roboter/ Verzeichnisdienst
Acoon	87 Mio. (weltweit)	09.2001	Roboter
AllesKlar	370 000 Sites in 10 000 Kategorien	05.2001	VD
AltaVista.com	550 Mio.	09.2001	Roboter
AltaVista.de	(82 Mio.)	11.2001	Roboter
Bellnet	400 000 Sites; Zahl der Kategorien nicht bekannt	11.2000	VD

<sup>3</sup> Verzeichnisdienste versprechen ohnehin Relevanz anstelle von Menge, die vergleichsweise kleineren Datenbestände werden hier nicht unbedingt als Nachteil angesehen.

### 3 Wissenswertes zu Suchmaschinen

Suchdienst	Index	Datum	Roboter/ Verzeichnis- dienst
Dino	275 000 Sites in 50 000 Katego- rien , 20 Sachge- biete	08.2000	VD
Euroseek	50 Mio.	08.2000	Roboter
AllTheWeb (FAST Seach & Transfer ASA)	625 Mio. basie- rend auf 1,5 Milli- arden erfasster Seiten	11.2001	Roboter
Fireball	8,2 Mio.	01.2001	Roboter
Google	2 Milliarden	12.2001	Roboter
HotBot (Inktomi)	ca. 600 Mio. basie- rend auf 1,5 Milli- arden erfasster Seiten	09.2001	Roboter
Infoseek (Deutschland)	42 Mio.	09.2000	Roboter
LookSmart	2,0 Mio. Sites in 200 000 Katego- rien	10.2000	VD
Lycos (weltweit)	250 Mio.	09.2000	Roboter
Northern Light	354 Mio.	05.2001	Roboter
ODP (Open Direc- tory Project)	2,839 Mio. Sites in 402 157 Katego- rien	09.2001	VD
Qualigo	15 Mio.	05.2001	Roboter
Scoutmaster	6,3 Mio.	09.2001	Roboter
Sharelook	182 000 Sites in 36 000 Rubriken, 12 Hauptkatego- rien	08.2000	VD
Speedfind	12 Mio.	08.2000	Roboter
Teoma	100 Mio.	09.2001	Roboter

### Indexgrößen wichtiger Suchmaschinen

Suchdienst	Index	Datum	Roboter/ Verzeichnis- dienst
Voila (weltweit)	145 Mio. einschl. Bilder, Videos	08.2000	Roboter
WiseNut	850 Mio (1,495 Milliarden Links)	09.2001	Roboter
WEB.DE	459 843 Sites in über 36 650 Kate- gorien	09.2001	VD

Halten Sie sich auf dem Laufenden, was es Neues im Suchmaschinen-geschäft gibt, und stellen Sie sicher, dass Sie wichtige Entwicklungen nicht verpassen. Es kann sein, dass zwei Jahre nichts passiert, dann kommt ein neues Angebot mit für den Benutzer überzeugenden Argumenten, und innerhalb kürzester Zeit ändern viele Benutzer ihre Präferenzen.



Es gibt noch weitere Entwicklungen, die zu beachten sind:

- Die Anzahl an echten Suchmaschinen-Indizes ist kleiner als die Menge von Anbietern des Dienstes für Benutzer. Viele Anbieter von Suchdiensten haben gar keine eigene Maschine, sondern nutzen die eines anderen Unternehmens, des eigentlichen Betreibers. Man nennt solch eine Zusammenarbeit in Anlehnung an ähnliche Verfahren bei Werbeanzeigen »Syndicating«.

Wenn man so will, kann jeder Anbieter eines solchen Suchdienstes sein. Suchmaschinen sind daran interessiert, aus möglichst vielen Webseiten genutzt und dadurch indirekt von ihnen empfohlen zu werden. Deswegen bieten viele die Möglichkeit an, ein kleines Suchformular auf die eigene Website zu platzieren, von dem aus eine Suche bei der Suchmaschine ausgelöst werden kann.

Unternehmen wie Yahoo oder MSN machen es grundsätzlich nicht anders, allerdings steht dahinter eine kommerzielle Vereinbarung. In Kapitel 8 finden Sie eine Übersicht über wichtige Suchmaschinen und sehen, wer dabei auf fremde Datenbestände zurückgreift.

- ▶ Einige Suchmaschinenbetreiber sind recht erfolgreich damit, die begehrten vorderen Plätze der Ergebnisseiten zu verkaufen oder sogar zu versteigern, zum Beispiel *www.overture.com*.

Speziell beim Versteigern von vorderen Eintragsplätzen ist die Hoffnung dieser Unternehmen, dass sich Wettbewerber im Bemühen um die beste Platzierung gegenseitig zu immer höheren Platzierungsgebühren steigern können.

Dieses Geschäftsmodell – Ranking gegen Geld – kommt nicht bei allen Unternehmen schlecht an:

Große Unternehmen müssen oft feststellen, dass sie bei den Ergebnissen weit hinter kleinen Anbietern stehen, die aktiver und einfallreicher waren, die besser wussten, worauf es ankommt, wenn man nach vorne kommen will, oder die einfach durch einen Zufall dort gelandet sind. Unternehmen, die bereit sind, für Platzierungen zu investieren, werden deshalb käufliches (oder sogar zu ersteigerndes) Ranking begrüßen.

Google geht einen interessanten anderen Weg: Im Rahmen der AdWord-Kampagne kann man günstig einen eher schlichten Werbeplatz neben der Ergebnisliste buchen, der Benutzern gezeigt wird, die bestimmte Suchwörter eingegeben haben, maximal zwei Anbieter je Ergebnisliste haben zusätzlich die teurere Möglichkeit, sich oberhalb der Ergebnisliste, aber optisch klar von dieser getrennt, zu platzieren. Wenn man akzeptiert, dass auch ein Anbieter eines Suchdienstes ein funktionierendes Geschäftsmodell braucht, welches ihm das wirtschaftliche Überleben sichert, ist ein solches Vorgehen sicher zulässig. Auf jeden Fall ist die Trennung von Suchergebnis (quasi dem redaktionellen Teil) und Werbung zu begrüßen.

Web.de macht es genau umgekehrt: Es gibt keine freie Anmeldung mehr, stattdessen muss man, um in das Verzeichnis aufgenommen zu werden, einen jährlichen Obolus zahlen, der schon für einen schlichten Eintrag bei 24 Euro liegt. In den Statistiken der Web-Beobachter, wie zum Beispiel *etracker.de*, ist Web.de nicht unter den großen Erzeugern von Verkehr zu den Websites, man sollte sich also überlegen, ob es den Einsatz lohnt.

- ▶ Viele Suchmaschinen haben Schwierigkeiten, profitabel zu sein. Ein Beispiel ist *Excite.de*, das Ende 2001 den Betrieb einstellen musste.

### *Wichtig: Nutzerverhalten einschätzen*

Es ist zu erwarten, dass sich die Zahl der Betreiber von Suchmaschinen in Zukunft nicht nennenswert steigern wird, selbst wenn neue Anbieter mit immer raffinierteren Ideen hinzukommen, und dass der kommerzielle Anreiz, Ranking zu verkaufen, weitere Anbieter dazu bringen wird, das Konzept zu übernehmen. Andererseits wird der wirkliche Erfolg einer Suchmaschine eher über die Zahl ihrer Nutzer definiert, und ob eine Suchmaschine, die empfunden wird wie ein Anzeigenblatt, noch genauso häufig aufgerufen wird, ist fraglich. Wenn dann die Nutzer ausbleiben, steht das Geschäftsmodell allgemein in Frage.

## **3.4 Wichtig: Nutzerverhalten einschätzen**

Die meisten Nutzer sind eher träge: Sie entscheiden sich einmal für eine Suchmaschine, und verwenden dann überwiegend diese für einen Zeitraum von vielen Monaten oder Jahren. Manche kennen alle Tricks der Profisuche, wie »und«, »oder«- oder »nicht«-Verknüpfungen, die bei einer anderen Suchmaschine anders funktionieren und erst neu gelernt werden müssen (siehe Kapitel 4). Außerdem sind sie überzeugt, dass der Mix aus Reaktionszeit, sowie Umfang, Relevanz und Aktualität der Ergebnisse für sie richtig ist. Auch Stilfragen sind vielen Benutzern wichtig, wie die klare Trennung von redaktionellem Teil – also den Suchergebnissen und eventuellen weiteren Inhalten, wie Nachrichten – auf der einen Seite und Werbeeinblendungen auf der anderen. Auch Pop-up-Fenster, die sich ungewollt beim Öffnen der Seite auftun, tragen nicht gerade zur Kundenzufriedenheit bei. Es ist sicher kein Zufall, dass die wirklich erfolgreichen Sites ohne sie auskommen.

Möglicherweise hat die für Sie interessante Benutzergruppe hinsichtlich Suchmaschinen ganz spezielle Vorlieben, die Sie kennen sollten? Möglicherweise ändern sich diese gerade jetzt. Vielleicht werden in naher Zukunft neue technische oder kommerzielle Angebote auf den Markt kommen, die von der von Ihnen anvisierten Zielgruppe dankbar angenommen werden?

Sie werden nicht umhin kommen, den Suchmaschinenmarkt laufend weiter zu verfolgen. Das Wissen aus diesem Buch mag Ihnen dabei helfen, dies mit einem kritischen Blick für das wirklich Wesentliche und für die Zusammenhänge zu tun.

