

Auf Beutezug im Internet ...

Die Anzahl der Websites im Internet wurde schon für das Jahr 2001 auf etwa 9 Milliarden geschätzt – eine extrem hohe Zahl. Etwa ein Drittel dieser Websites war über Suchmaschinen zugänglich – der Rest ist zurzeit nur für Eingeweihte oder per Zufallstreffer zugänglich.

Um zumindest an den »bekannteren« Teil des Internets zu gelangen, wollen wir uns im ersten Teil mit der Bedienung der Suchmaschinen beschäftigen: Welche Unterschiede gibt es bei den verschiedenen Suchmaschinen und wie benutzt man sie sinnvoll. Im zweiten Teil geht es dann darum, wie man den »Fang« dauerhaft auf seine Festplatte bannt.

Eine kleine Vorbemerkung: Das Internet ist ein extrem schnelllebiges Medium – eine heute vorhandene Website kann morgen schon für immer vom Netz sein, eine heute empfohlene kommerzielle Suchmaschine kann morgen schon pleite sein – oder der engagierte Privatbetreiber hat irgendwann die Lust verloren ...

1. Suchen und finden

Um etwas zu finden, gibt es zwei Möglichkeiten: Entweder man hält Ordnung oder man muss lange suchen. Hat man eine sinnvolle Ordnung eingehalten, ist man schnell am Ziel, muss man im Chaos suchen, würde man diese Aufgabe gerne jemand übertragen, dem man einfach nur sagt »Das was ich suche sieht so und so aus, guck 'mal, ob Du es findest ...«. Es ist klar, dass der Suchende dann auch mit einer Menge Müll ankommt, welcher nur ungefähr dem gesuchten Gegenstand entspricht.

Aber genau nach diesen beiden Ansätzen funktioniert die Suche im Internet: Zum einen gibt es *Menschen(!)*, die damit beschäftigt sind, die Websites des Internets in einem **Katalog** (z.B. Yahoo oder Lycos) so zu sortieren, so dass man einen einfachen Zugriff zu Websites bekommt. Zum anderen gibt es »robots« (auf Computerfarmen installierte Suchprogramme), die einfach alle Wörter einer Website durchlesen (**Volltextsuchmaschinen** wie Google, Altavista, Fireball), sich diese in einer Liste merken und auf die Anfrage »Wo kommt überall ›Buttermilch‹ als Begriff vor«, sofort alle infrage kommenden Websites präsentieren.

Als dritte Art gibt es noch »Schmarotzer«, die eine Anfrage an andere Suchmaschinen weiterleiten, und deren Ergebnisse dann in einer eigenen Liste präsentieren: **Metasuchmaschinen** (MetaGer, Vivisimo, ...)



„Go ask your search engine“

Volltextsuchmaschinen

durchsuchen das Netz maschinell, es gibt *keine* Redaktion, die Abdeckung des Netzes beträgt etwa 30%. Volltextsuchmaschinen nehmen mit Ausnahme von Stoppwörtern (»die«, »und«, »the«, ...) alle auf einer Website vorkommenden Wörter auf – dies gewissermaßen »ohne Sinn und Verstand«.

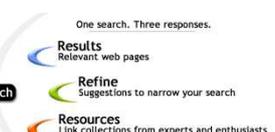
- **www.google.de** kannte z.B. im April 2002 zwei Milliarden der etwa neun Milliarden Websites des Internets. Google benutzt ein Verweis-Ranking, d.h. je mehr Verweise im Netz auf eine bestimmte Website zeigen, desto weiter oben erscheint das Suchergebnis in der Trefferliste. Auch wenn der gesuchte Begriff im Titel oder gar Domainnamen vorkommt, landet solch eine Website weit oben auf den Trefferlisten von google.
- **www.teoma.com**
- **www.wisenut.com**
- **www.altavista.de / com**
- **http://search.msn.de/**
- **www.fireball.de** (incl. »live-suche« – was suchen andere gerade?)



Suche: Das Web Seiten auf Deutsch Seiten aus Deutschland



Search with Authority™ Find this phrase



Kataloge

sind redaktionell betreute Schlagwortverzeichnisse – durch die »menschliche« Betreuung vom Prinzip her »hochwertiger« als maschinell erzeugte Volltextsuchmaschinen.

In meist etwas mehr als zehn Hauptkategorien (Rubriken, Channels, Portale, ...) eingeteilt, hangelt man sich durch eine hierarchisch aufgebaute Baumstruktur in das Gebiet, das einen interessiert und startet dort die Suche (und nicht gleich auf der Startseite!). Teilweise wird man bei Nichterfolg automatisch an eine Volltextsuchmaschine weitergereicht (»Suche in Websites«).

Wegen des Personalaufwandes oft nur ein kleiner Ausschnitt des WWW, zeitverzögerte Aufnahme von Seiten, oft auf bestimmte Sprachen beschränkt.

- <http://dmoz.org/World/Deutsch/> (Open Directory Project, einen Teil dieses Katalogs findet man auch, wenn man auf der Startseite von google auf den Link »Verzeichnis« geht)
- www.yahoo.de
- www.allesklar.de
- www.lycos.de
- <http://rubriken.fireball.de/>
- <http://search.msn.de/>
- <http://bubl.ac.uk/link> – akademisch orientiert
- www.web.de – Bild-Zeitungs-Niveau

} gehören zusammen ...



Kategorien

- Bildung & Ausbildung**
Uni/FH, Schulen, Bewerbung...
- Computer & Internet**
Internet, Hardware, Software, Chat...
- Geisteswissenschaften**
Psychologie, Geschichte, Sprachen...
- Gesellschaft & Soziales**
Essen & Trinken, Sexualität, Gay...
- Gesundheit**
Medizin, Krankheiten, Psyche, Diät...
- Handel & Wirtschaft**
Firmen, Finanzen, Jobs, Immobilien...
- Kunst & Kultur**
Malerei, Literatur, Theater, Mode...
- Nachrichten & Medien**
Top Themen, TV, Zeitschriften, Zeitungen...
- Nachschlagewerke**
Lexika, Wörterbücher, Telefonnummern...
- Naturwiss. & Technik**
Tiere, Pflanzen, Astro, Handy und SMS...
- Sport & Freizeit**
Autos, F1, Fußball, Ski, Spiele, Reisen...
- Staat & Politik**
Politik, Behörden, Militär, Recht, Steuern...
- Städte & Länder**
Dt. Städte, Länder, Europa, Landkarten...
- Unterhaltung**
Cooler Links, Musik, MP3, Humor, Kino...

Kategorien



- Arbeit & Beruf**
Arbeitsämter, Gewerkschaften, Jobbörsen & Stellensuche, Arbeitnehmer, ...
- Auto & Verkehr**
Automarken A-Z, Verkehrsinformationen, Motorsport, Fahrzeughandel, ...
- Bildung & Wissenschaft**
Schule, Unterricht & Materialien, Forschung, Studium & Universitäten, ...
- Computer & Internet**
Computerspiele, Hardware, Internet, Software, Tipps & Hilfen, ...
- Finanzen & Wirtschaft**
Banken & Sparkassen, Börse, Investieren & Anlegen, Versicherungen, ...
- Firmen & Behörden**
Dienstleistungen, Handel, Herstellung & Produktion, Verzeichnisse & Infos, ...
- Freizeit & Lifestyle**
Essen & Trinken, Lesen, Kino & Filme, Mode & Schönheit, Veranstaltungen, ...
- Gesellschaft & Politik**
Staat, Religion & Kirche, Recht & Gesetz, Hilfsorganisationen, Parteien, Soziales, ...
- Gesundheit & Medizin**
Krankenhäuser, Krankheiten, Apotheken, Ärzte, Schwangerschaft & Geburt, ...
- Immobilien & Wohnen**
Immobilienbanken, Mieter & Vermieter, Bauen, Makler & Vermittler, ...
- Kunst & Kultur**
Film & Video, Musik, Literatur & Bücher, Museen & Galerien, Künstler, ...
- Medien & Informationen**
Fernsehen, Film & Video, Gut zu wissen, Presse & Nachrichten, Radio, ...
- Menschen & Gruppen**
Persönliche Homepages, Familie, Jugendliche & Kinder, Senioren, ...
- Reisen & Tourismus**
Unterkünfte, Länder- & Stadtinfos, Ausflugsziele, Last Minute, Reisebüros, ...
- Sport & Fitness**
Allg. Sportvereine, Fußball, Tanzsport, Wettkämpfe & Turniere, ...
- Umwelt & Technik**
Energie, Ingenieure, Klima, Lebensräume, Umweltorganisationen, ...

Kategorien



- Schlagzeilen, News**
- Auskunft, Telefon**
- Auto, Motorrad**
- Beruf, Karriere**
- Boulevard**
- Computer, Download**
- Erotik**
- Film, Kino, DVD**
- Finanzen, Börse**
- Handy, Logos, Töne**
- Internet**
- Kultur, Kunst, Literatur**
- Musik, MP3, Radio**
- Reise, Urlaub**
- Sport, Fußball, F1**
- Städte, Regional**
- Wellness, Gesundheit**
- Wissenschaft**
- Wohnen, Immobilien**

Metasuchmaschinen

beauftragen eine Reihe von Suchmaschinen auf einmal (bis zu 100!) und sortieren aus deren Treffern in der Regel Dubletten heraus .

- **MetaGer** (<http://meta.rrzn.uni-hannover.de>) BOOLE, einstellbare Suchzeit
- **www.vivísimo.com** BOOLE, einstellbare Suchzeit, Kategorienbildung
- **www.nettz.de** BOOLE
- **www.search.com** über 100 Suchdienste, BOOLE
- **www.profusion.com** BOOLE, einstellbare Suchzeit
- **www.suchen.com**



Wie fragt man richtig ?S

Folgende Operatoren aus der Logik (»BOOLEsche Operatoren«) helfen bei den meisten Suchmaschinen. Genaueres findet man auf den Hilfe-Seiten der benutzten Suchmaschine – oft muss man auch erst auf die »Profi-Suche«, »Advanced Search«, »Erweiterte Suche«-Seiten der Suchmaschine gehen.

Zeichen	Bedeutung	Beispieleingabefindet
*	Platzhalter, Trunkierung	Schul*	»Schule«, »Schulen«, »Schulentwicklung«, »Schulstress«, ... und schon hat man 1.078.459 links ...
		Sch*lervertretung	findet Schülervertretung, Schuelervertretung, Schülervvertretung (HTML)
AND UND + &	Beide Begriffe müssen <i>irgendwo(!)</i> auf der Website vorkommen (meist automatisch von den Suchmaschinen ergänzt, falls zwei Begriffe eingegeben werden)	Lehrplan+Physik	»... hält sich kaum an den Lehrplan ...mein Studienwunsch ist Physik ...«
" "	Phrasensuche Der Begriff zwischen den Hochkommata muss – einschließlich Leerzeichen – exakt so auf der Website vorkommen	"Lehrplan Physik"	»... offizielle Lehrplan Physik für bayerische Gymnasien... «
OR ODER	Einer von beiden Begriffen muss auf der Website vorkommen		
NOT – !	Negation – muss <u>direkt</u> vor dem Begriff stehen »Hafen-Fähre« ist etwas anderes als »Hafen –Fähre«	Kohl NOT Kanzler	»... ein leckerer Salat mit Kohl ...«
NEAR ~	Näherungssuche	»Kohl NEAR Kanzler« <i>bringt andere Ergebnisse als</i> »Kohl NEAR Kartoffeln«	
()	Logische Klammern		

Bsp. »Kanzler NEAR Kohl AND Schäuble NOT Schwarzgeld«

Klassische Nachschlagewerke befragen

Eigentlich ist es sinnvoller, mit den jeweiligen CDROM-Versionen der unten genannten Nachschlagewerke zu arbeiten, die (im Vergleich zu den mehrbändigen Buchversionen) recht preiswert sind. Aber nicht immer lohnt sich die Anschaffung – oder man braucht unterwegs eine Erklärung.

Enzyklopädien

- <http://encarta.msn.de> enthält knapp die Hälfte der Microsoft Encarta-CD-ROM im Online-Angebot. Die Encarta-Artikel haben einen guten Ruf, wenn man dort fündig wird, braucht meist nicht mehr weitersuchen. Achtung : Microsoft lässt nur den Internet Explorer an seine Daten ran – aber zum Glück kann sich Opera ja als IE 5.0 ausgeben ...
- **www.britannica.com** enthält den gesamten Artikelbestand der Buchversion, allerdings jetzt nicht mehr umsonst ...
- **www.xrefer.com** ist für allgemeine Anfragen gut und kostenlos

Nachschlagewerke

- **www.wissen.de** (Bertelsmann Verlag) bietet den kompletten Inhalt von »Das große Bertelsmann Lexikon«, »Bertelsmann Universallexikon 2002« und »infoline 2002« kostenlos online an.
- **www.xipolis.net** Bibliographisches Institut (Brockhaus, Duden, Langenscheidt, Meyer) und Holtzbrinck (Rowohlt, Fischer, ..., Zeit, Tagesspiegel, Handelsblatt, ..., Spektrum, natur, Scientific American, ...) teilweise kostenlos, die besten Artikel kosten hier allerdings Geld ...
- **www.dictionary.com** Englischsprachige Wörterbücher und Thesaurus



Zum Abschluss: Suchstrategien

Faustregel 1: Je spezieller der zu suchende Begriff („Glasaugenherstellung“) ist, desto breiter muss die Suche angelegt werden (d.h. Metasuchmaschinen benutzen).

Umgekehrt versucht man es bei relativ allgemeinen Suchbegriffen zuerst in einem guten Katalog.

Faustregel 2: Man sollte sich nach einigen Tests auf ein paar »erste Anlaufstellen« festlegen.

Jede Volltextsuchmaschine hat ihre Besonderheiten z.B. bezüglich der Abfragesyntax. Wenn man diese beherrscht, ist dies schon die Hälfte der Miete. Ebenso kennt man irgendwann auch die Kategorien eines Kataloges so gut, dass man sich relativ schnell darin bewegen kann.

Faustregel 3: Tagsaktuelle Ereignisse kann bei allgemeinen Suchmaschinen und Katalogen nicht finden – sie haben einen Nachlauf von etwa 3 Wochen. Hier muss man die Zeitungs-Suchmaschinen benutzen.

2. Finden und behalten

Ein einzelnes Bild »grabben«

Opera	mit rechter Maustaste Bild anklicken: »Bild speichern ...«
Internet Explorer	mit rechter Maustaste Bild anklicken: »Bild speichern unter ...«
Mozilla	mit rechter Maustaste Bild anklicken: »Grafik speichern unter ...«

Speichern von ganzen Seiten

Opera	»Datei – Speichern ...« (ohne Bilder) »Datei – Mit Bildern speichern ...« (mit Bildern)
Internet Explorer	»Datei – Speichern unter ...« (mit Bildern ist voreingestellt, sonst »Dateityp« ändern)
Mozilla	»Datei – Speichern unter ...« (mit Bildern ist voreingestellt, sonst »Dateityp« ändern)

Das Laden einzelner Dateien (»Energiesparen.pdf«)

Falls das Anklicken eines Download-Links zu einem Fenster mit wirren Zeichen im Browser führt, muss man mit dem »Zurück«-Button auf die vorherige Seite zurückgehen, und den Link mit der rechten Maustaste anklicken. Aus dem Kontextmenü »Ziel speichern unter ...« wählen.

Speichern vollständiger Websites mit einem Ripper/Offline-Browser

Websites kommen und gehen. Gestern war noch eine attraktive Seite im Netz, heute schon ist sie durch eine neue Seite ersetzt worden. Auch möchte man so manche Website mit ihren links in Ruhe lesen, ohne dass der Gebührenzähler läuft ...

Websites kann man komfortabel mit sogenannten Rippern speichern. Einmal angeklickte Seite bleiben auf der Festplatte, man kann sogar die Software beauftragen, alle links auf der Seite zu verfolgen und ebenfalls herunterzuladen und zu speichern. Auf diese Art kann man Informationen einmal »online« zusammensammeln, man surft in Windeseile auf der eigenen Festplatte – und es fallen keine Gebühren an.

Wysigot (www.wysigot.com) ist intuitiv zu bedienen und komfortabel. Die Freeware-Version ist natürlich nicht ganz so komfortabel wie die Vollversion.

Hyperfolio: (www.schule-online.de/hyperfolio) »... *HyperFolio ist ein völlig neuartiges Werkzeug zum Ordnen und Organisieren von Informationen aus dem Internet. Mit HyperFolio können Sie Medienobjekte von Ihrem Browser oder einer anderen Anwendung direkt zum HyperFolio-Sammelordner ziehen und dort ablegen. Anders als Lesezeichen -- mit denen Sie nur eine Internetseite auf einmal aufschlagen können -- ermöglicht HyperFolio es Ihnen, Texte, Bilder, Audio- und Videodateien direkt von Ihrem Browser aus zu sammeln, zu katalogisieren und zu kommentieren ...*« 30-Tage-Trialversion, Vollversion für 9,90 €.