

X1 Google – Ein Pionier



Die Suchmaschine Google ist in kürzester Zeit vom Geheimtipp zur am meisten verwendeten allgemeinen Suchmaschine im Internet aufgestiegen. Deshalb verdient Google ein eigenes Kapitel. Google gibt es in den Varianten www.google.com, www.google.de und auch www.google.ch.

In diesem Posten wird anhand von verschiedenen Originalartikeln der Werdegang von Google aufgezeigt.

- | |
|--|
| Artikel 1: „ Warum Google? “
Selbstwerbung des Suchdienstes |
| Artikel 2: „ Suchmaschine Google ist der Retter genervter Surfer “
Berliner Morgenpost, 3.3.2001 |
| Artikel 3: „ Die Erfolgsstory von Google “
World online |

Artikel 1:
Selbstwerbung des Suchdienstes

Warum Google?

Weil Google schnell die relevantesten Suchergebnisse liefert!

Mit Hilfe der ausgefeiltesten wegweisenden Technologien findet Google schnell die richtigen Ergebnisse für jede Suchanfrage. Google gibt relevante Ergebnisse zurück, weil die gefundenen Web-Sites automatisch anhand der Link-Struktur des Internets bewertet werden.

Der überwältigende Umfang der Informationen im Web erfordert einen hervorragenden Suchdienst, der diese Informationen zugänglich macht und den Benutzern erschließt. Ohne ein leistungsfähiges Suchwerkzeug kann die Suche nach einer bestimmten Web-Site so schwierig sein wie die Suche nach einem Buch in einer Bibliothek ohne Katalog, in der die Bücher ohne Ordnungsprinzip willkürlich in die Regale gestellt werden.

Google bringt Ordnung ins Web

Google ordnet das Informationschaos. Google ist ein Suchdienst, wie er sein soll: kein redigiertes, eingeschränktes Verzeichnis oder eine Liste von Ergebnissen, die an den Meistbietenden versteigert wurden, sondern eine durchdachte Methode, das Internet anhand seiner eigenen Struktur zu organisieren.

Google revolutioniert die Web-Suche mit seiner zum Patent angemeldeten PageRankTM-Technologie. PageRank nutzt die Struktur des Webs, die durch die Art definiert wird, wie eine Web-Seite direkt, unmittelbar und ohne Drittparteien mit anderen Web-Seiten verknüpft werden kann. In einer gewissen Weise demokratisiert diese Link-Struktur das Internet auf automatischem Wege. Sie enthierarchisiert die Informationen und ermöglicht den ungehinderten Fluss der Informationen und Ideen zwischen den einzelnen Sites.

So funktioniert PageRank

PageRank nutzt die umfassende Link-Struktur des Webs als

Organisationskriterium und profitiert so von dessen einzigartiger demokratischer Beschaffenheit. Im wesentlichen interpretiert Google einen Link von Seite A zu Seite B als eine Stimme *von* Seite A *für* Seite B. Google beurteilt die Wichtigkeit einer Seite nach der Anzahl der abgegebenen Stimmen. Google berücksichtigt jedoch nicht nur die Anzahl der Stimmen bzw. Links, sondern analysiert auch die Seite, von der die Stimme ausgeht. Stimmen von Seiten, die selbst als "wichtig" eingestuft werden, haben eine größere Bedeutung bei der Bewertung der Wichtigkeit anderer Seiten.

Diese wichtigen, qualitativ hochwertigen Ergebnisse erhalten einen höheren PageRank-Wert. Dieser bestimmt, an welcher Stelle in der Liste der Ergebnisse die zurückgegebene Seite erscheint. Auf diese Weise ist der PageRank-Wert ein Indikator für die allgemeine Wichtigkeit einer Seite, der nicht von einer bestimmten Suchanfrage abhängt, sondern eine Eigenschaft der Seite selbst darstellt. Er wird aufgrund von Daten aus dem Web bestimmt, die Google mittels komplexer Algorithmen, die die Link-Struktur bewerten, analysiert.

Natürlich ist eine wichtige Seite für Sie wertlos, wenn Sie nicht Ihrer Suchanfrage entspricht. Daher wendet Google ausgeklügelte Textanalysetechniken an, um sowohl wichtige als auch für Ihre Suchanfrage relevante Seiten zu finden. Wenn Google eine Seite analysiert, wird beispielsweise auch berücksichtigt, welche Informationen zu dieser Seite auf den Seiten enthalten sind, die auf die analysierte Seite verweisen.

Integrität

Die komplexen, automatisierten Suchtechniken von Google schließen jeden menschlichen Eingriff aus. Im Gegensatz zu anderen Suchmaschinen verhindert die Struktur von Google, dass jemand einen höheren PageRank-Wert erwerben kann oder die Ergebnisse auf andere Weise kommerziell verfälscht. Eine Google

Suche ist eine ehrliche und objektive Methode, auf einfache Weise qualitativ hochwertige Web-Sites zu finden.

Google gibt nur die Seiten zurück, welche die eingegebenen Suchbegriffe enthalten.

Im Gegensatz zu vielen anderen Suchmaschinen gibt Google nur solche Ergebnisse zurück, die entweder im Seitentext oder in den Namen der Links, die auf die Seite verweisen, alle Suchbegriffe enthalten. Sie brauchen sich nie wieder mit Ergebnissen herumzürgeren, die gar nichts mit den von Ihnen eingegebenen Suchbegriffen zu tun haben.

Google beachtet die Position Ihrer Suchbegriffe auf einer Seite.

Die Ergebnisse von Google enthalten nicht nur alle eingegebenen Suchbegriffe; Google analysiert auch, wie nahe diese Begriffe auf einer Seite beieinander stehen. Im Gegensatz zu vielen anderen Suchmaschinen bewertet Google die Relevanz der Ergebnisse auch anhand der Nähe der Suchbegriffe. Wir bewerten Ergebnisse, bei denen die Suchergebnisse nahe beieinander liegen, höher, damit Sie sich nicht durch eine Unmenge von irrelevanten Ergebnissen arbeiten müssen.

Google bietet eine aussagekräftige Vorschau für jedes Ergebnis.

Statt feststehender Zusammenfassungen zeigt Google in den Suchergebnissen einen Auszug aus dem Text an, der auf Ihre Suchanfrage abgestimmt ist, und in dem Ihre Suchbegriffe hervorgehoben sind. Dies erspart Ihnen die Zeit und den Ärger, irrelevante Web-Seiten zu laden.

Mit Google haben Sie Glück!

Google ist unübertroffen darin, das richtige Ergebnis für eine allgemeine Suchanfrage (wenn der Suchbegriff zum Beispiel ein Firmenname ist) an erster Stelle zurückzugeben. Wir vertrauen dieser Funktion so sehr, dass wir die Schaltfläche

"Auf gut Glück™!" eingerichtet haben, die Sie direkt zur Web-Site des ersten Suchergebnisses führt. Mit der Funktion "Auf gut Glück!" von Google gelangen Sie so schnell wie möglich an die gewünschten Informationen.

Google speichert Web-Seiten im Cache.

Google speichert in seinem Cache zahlreiche Kopien von Web-Seiten, auf die Sie bei einem zeitweiligen Ausfall des Servers der Web-Seite zurückgreifen können. Wenn der Server nicht verfügbar ist, kann die Kopie der benötigten Seite im Cache von Google Ihr Retter sein. Häufig können Sie auf das Material im Cache schneller zugreifen als über den Link zur eigentlichen Seite; die auf diese Weise erhaltenen Informationen sind jedoch möglicherweise nicht auf dem neuesten Stand. In den meisten Fällen ist das aber immer noch besser als ein "404 Not Found"-Fehler!

Artikel 2:

Berliner Morgenpost, 3.3.2001

Suchmaschine Google ist der Retter genervter Surfer

Von Christiane Schulzki-Haddouti

Eine kleine, unscheinbare Suchmaschine hat sich in wenigen Monaten vom Geheimtipp zur größten und intelligentesten Suchmaschine im Netz entwickelt: Google. Dabei liest sich die Geschichte der Firma wie die berühmte Amerika-Story vom Aufstieg des Tellerwäschers zum Millionär.

Auf den ersten Blick sieht Google wie ein Suchdienst irgendeiner Universität aus. Bis auf ein einsames Eingabefeld und das Logo ist nichts zu sehen. Keine Kataloge, keine Nachrichten, kein Übersetzungsdienst, keine Werbung - kein Portal. Google kennt nur eins: das Suchen von Informationen. Und, erstaunlich für den Erstnutzer, oft stehen die relevanten Ergebnisse an erster Stelle. Der Trick: Google bewertet die Webseiten als wichtig, auf die viele andere Webseiten

verweisen. Die am häufigsten «Verlinkten» stehen ganz oben.

Entwickelt wurde Google vom Computer-Ingenieur Lawrence Page und dem Mathematiker Sergey Brin. Sie lernten sich an der Stanford University kennen, wo sie «einen Prototypen für eine umfassende Suchmaschine» entwickelten. Heute ist Google ein florierendes Unternehmen. Etwa 40 der 200 Mitarbeiter haben einen Dokortitel, die Hälfte ist im Software-Engineering tätig. Ohne Werbung, nur durch Mund- oder besser Netzpropaganda entwickelte sich Google rasch zum Geheimtipp. Im Januar 2001 lag die Firma mit 8,8 Millionen Besuchern bereits auf Platz 25 der US-Websiteliste von Media-Metrix und rangiert unter den beliebtesten Suchmaschinen auf Rang fünf. [...].

Vor einem Jahr war der Google-Index noch 150 Millionen Seiten lang. Im Winter 2000 holte er auch im Indizierungswettbewerb gegen Konkurrenten wie Altavista und Inktomi auf. Jetzt erfasst sie die meisten Websites überhaupt: Im November wurden 602 Millionen Webseiten von ihr laut Searchenginewatch.com indiziert. Altavista kannte hingegen gerade einmal 350 Millionen. Da Google jedoch auch Links berücksichtigt, erfasst es auch Seiten, die es selber nie besucht hat. Damit deckt es rund 1,3 Milliarden Seiten ab - im Juli 2000 wurde der Umfang des Webs auf zwei Milliarden Seiten geschätzt. [...].

Das Google-Archiv belegt komprimiert heute mehrere hundert Terabyte Festplattenplatz. Der älteste Eintrag im Index ist rund 30 Tage alt. Praktisch: Neben dem Index speichert Google auch die Webseiten selbst, um bei den Treffern die entsprechenden Suchbegriffe in einer Art Kurzzusammenfassung darstellen zu können.

Hinter Googles schlichter Fassade steckt ein gewaltiger Hardware-Park: Rund 7000 Computer sind auf fünf Standorte an der West- und Ostküste der USA verteilt. Sie bilden die Hardware-Ausstattung von Google, laufen unter dem freien

Betriebssystem Linux und werden von einer kleinen Firma extra für Google assimiliert. Hardware und Software sind für die Parallelverarbeitung ausgelegt. 70 Millionen Abfragen muss die Suchmaschine inzwischen jeden Tag verkraften. Anders als die Portale Altavista.com oder Yahoo.com fungiert der Suchdienst selbst als riesiges Werbeportal für die Firma Google - und sonst für niemanden. Google lizenziert seinen Suchdienst an andere Firmen, wie zum Beispiel Cisco, die ihre eigenen Websites damit erschließen. Aber auch Internetportale haben sich bereits zu einer Google-Lizenz durchgerungen. Im letzten Sommer entschloss sich sogar das amerikanische Yahoo!-Portal, seine Suchmaschine mit Google-Technik auszurüsten - und sich von Inktomi zu verabschieden. Für Farzad Nazem, Chef der Entwicklungsabteilung bei Yahoo!, war klar: Google bietet derzeit die beste Suchtechnologie. Für manche Nutzer reduziert sich das Suchen derzeit auf zwei Klicks: Bei Google suchen, bei Altavista-Babelfish übersetzen.

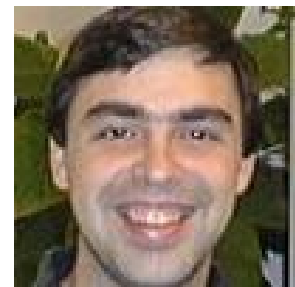
Das Geheimnis des Erfolgs ist, dass die Google-Entwickler Brin und Page nicht auf Schlüsselwörter (Altavista) oder persönliches Ranking (Yahoo!) setzen, sondern auf Links. Sie schufen Karten, die die Vernetzung über Hyperlinks wiedergaben. Auf diese Weise entsteht ein Page-Ranking, vergleichbar dem Bewertungssystem unter Wissenschaftlern: Je häufiger seine Artikel und Werke zitiert werden, umso angesehener wird er. Und: Die Zitate sind auch ein Hinweis auf die Qualität dieser Werke. Gleiches gilt im Netz. Das Page-Ranking von Google widerspiegelt aber auch das Webverhalten: Surfen ist oft zufälliges Springen von einer zur anderen Seiten auf der Flucht vor Langeweile. Kombiniert mit anderen Methoden macht das die Stärke von Google aus. So wird im Ranking auch berücksichtigt, welche Seiten zu einer anderen Seite verweisen. Dabei bestimmt die Ranghöhe dieser Seiten wiederum die Ranghöhe der anderen Seiten. Wenn eine

Seite zum Beispiel nur einen einzigen Link von Yahoo! erhält, liegt sie bei den Google-Suchergebnissen weiter vorn. Und: Die Suche erfasst nicht nur Text-Links, sondern auch Verweise auf Bilder und Programme. Google verfügt noch über weitere Features. Es merkt sich die Hits und generiert darüber Suchwahrscheinlichkeiten. Zudem gewichtet es auch visuelle Darstellungen: Wörter, die fett oder größer dargestellt sind, werden höher gewichtet als andere Wörter. Seit kurzem kann Google auch alle PDF-Dokumente durchsuchen und ist mit 13 Millionen PDF-Dateien sogar eine bessere Alternative als Adobes eigener Suchservice mit nur einer Million Dateien. Der Clou: Die PDF-Dateien sind bei Google als Text-Dateien abrufbar.

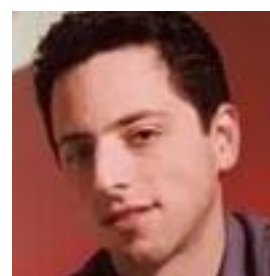
Artikel 3: World online

Die Erfolgsstory von Google

Die Suchmaschine Google wurde im Jahre 1998 an der Stanford Universität von zwei jungen Informatikern entwickelt und hatte Ende September 1999 ihren ersten Online-Auftritt.



Sergey Brin und Larry Page, so die Namen der beiden Google-Schöpfer, ist es gelungen, eine raffinierte Technologie zu entwickeln, mit der auf jede Suchanfrage so rasch und genau wie möglich eine Antwort gefunden wird.



Google sortiert die Suchergebnisse, indem es die Wichtigkeit jeder Webpage mit mathematischen Methoden bewertet. Das bedeutet einen Check von über 500 Millionen Variablen und 2 Milliarden Begriffen (!).

Mathematik und Genauigkeit

Die Besonderheit dieser neuen Technologie, die „Page Rank“ genannt wird, besteht in der Bewertung der Anzahl und der Qualität der Links, die sich auf eine Seite beziehen. Nachdem die inhaltliche Suche erfolgreich abgeschlossen ist, werden die Ergebnisse nach der Wichtigkeit der Webseiten geordnet, die auf die gefundenen Sites verweisen.

So findet sich z.B. eine Wirtschaftsanalyse eines vollkommen Unbekannten auf die von einer anderen unbekanntem Quelle verwiesen wird, am Ende der Ergebnisliste, während sich an der Spitze Artikel von bekannten Wirtschaftsgurus befinden, auf die direkt von der New York Times verwiesen wird. Google betrachtet die Links, die eine Webseite mit einer anderen verbinden als demokratisches Urteil, das von der verweisenden Seite über jene Seite getroffen wird, auf die verwiesen wird. Aus der Gesamtheit aller dieser „Bewertungen“ im Internet wird dann schliesslich die Wichtigkeit einer Internetseite berechnet. Auch die jeweiligen Seiten, auf denen sich ein Link befindet, werden auf dieselbe Art und Weise unter die Lupe genommen.

Es versteht sich also von selbst, dass nicht alle Seiten von gleicher Wichtigkeit sind: Das Geheimnis liegt in der Kontrolle der Ursprungsseiten und in der Zuordnung eines mehr oder weniger hohen „Page Rank“ je nach der Wichtigkeit der Webseite, die die Links enthält. Zudem bietet Google seinen Benutzern Suchergebnisse, die genau der Suchanfrage entsprechen. Anders als bei vielen anderen Suchmaschinen werden nämlich nur Webseiten angezeigt, die alle Schlüsselbegriffe enthalten, wobei überdies noch die Nähe der einzelnen Begriffe innerhalb des Dokuments analysiert wird. Die Gesamtheit dieser Überprüfungskriterien ermöglicht es Google, höchstpräzise Antworten zu liefern.

„I’m feeling lucky“

Daneben bietet Google noch eine ungewöhnliche Option, die es den Surfern erlaubt, sich Google blind anzuvertrauen und direkt auf das erste Link der Ergebnisliste zu springen. Diese aussergewöhnliche Option, die es den Usern erlaubt, so schnell wie möglich zu nützlichen Informationen zu kommen, wurde „I’m feeling lucky“ getauft und hat ihren grössten Nutzen bei häufig gestellten Anfragen, wie z.B. wenn es darum geht, die Homepage grosser Unternehmen oder berühmte Produkte zu finden.

29 Hundertstelsekunden

Die hohe Bearbeitungsgeschwindigkeit ist ein weiteres Markenzeichen von Google. Mit einer Durchschnittsdauer von 29 Hundertstelsekunden pro Anfrage zählt die Suchmaschine zu den schnellsten im ganzen Internet. Ende Juni konnte Google überdies die Erfassung von über einer Milliarde Webseiten vermelden (davon 560 Millionen vollständig) und damit Inktomi als Yahoo-Lieferant ablösen, was die Suche von Sites betrifft, die sich nicht in den Yahoo-Archiven befinden.

Klar und übersichtlich

Die Besonderheit, die an der Homepage von Google am meisten ins Auge sticht, ist ihre Einfachheit. Wer bis dato an die Unübersichtlichkeit anderer Suchmaschinen gewöhnt war, wird zweifelsohne positiv überrascht von der Klarheit und der spartanischen Eleganz von Google sein. Angesichts des Erfolgs von Google haben nun auch andere Suchmotoren damit begonnen, ihre Homepages nach dem minimalistischen Vorbild Googles zu gestalten.
