

Skript

D Bewerten von Suchergebnissen

Stand: Februar 2011

Achtung!

Skript ist nur zur Internen Verwendung im Seminar gedacht.

1 Grundsätzliches zur Bewertung von Suchergebnissen, Qualitätsmaßstäbe und Rankingmechanismen

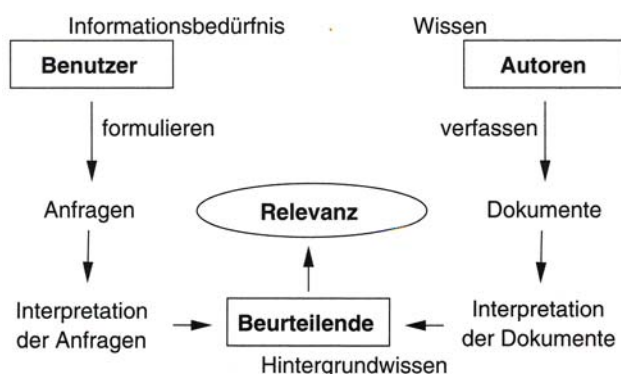
nach Hartmann et al. 2000, S. 30-48

1.1 Probleme bei der Informationssuche

Informationssuche bedeutet zum einen, das Finden von Informationen, die man noch nie zuvor gesehen hat. Zum anderen bedeutet es das Finden von Dokumenten, die die gesuchte Information *hoffentlich* enthalten.

Eine Suche kann also nur erfolgreich sein, wenn jemand ein Dokument mit der gewünschten Information verfasst hat, dieses zugänglich gemacht hat und dieses Dokument auch gefunden wird.

Die Schwierigkeiten werden deutlich: Nur was da ist, kann auch gefunden werden. Und: Wie bekommt die Suchmaschine heraus, welche Informationen gesucht werden, wie macht es diese Informationen dann zugänglich?



Grundsätzliches Problem der Informationssuche:

Auf der einen Seite steht der Benutzer mit einem Informationsbedürfnis: Dieses formuliert er in einer Anfrage als Menge von Suchbegriffen. Auf der anderen Seite

stehen Autoren mit einem bestimmten Wissen. Sie verfassen Dokumente, die auf diesen Informationen basieren.

Zwischen beiden steht in der Offline-Welt eine beurteilende Instanz, die sowohl die Anfrage als auch den Bestand an Dokumenten interpretiert. Sie besitzt Hintergrundwissen im gesuchten Feld und die Fähigkeit, Hintergrundwissen mit dem Wissen im Dokument zu verknüpfen. Deshalb kann sie gut entscheiden, welche Dokumente relevant für die Anfrage sind. Diese beurteilende Instanz ist zumeist entsprechend gebildet (Bibliothekar, Lehrer) und kennt zumeist auch das Wissensbedürfnis des Fragenden sehr genau. Zudem kann ein Mensch die komplexen Bedürfnisse von anderen Menschen sehr gut ‚übersetzen‘. In der Offline-Welt funktioniert Informationssuche durch INTERPRETATION sehr gut.

Beispiel: Ein Journalistik-Student fragt einen befreundeten Anglistik-Studenten nach Hintergrundmaterial zum diesjährigen Literaturnobelpreisträger.

Informationssuche mittels elektronischem Suchsystem

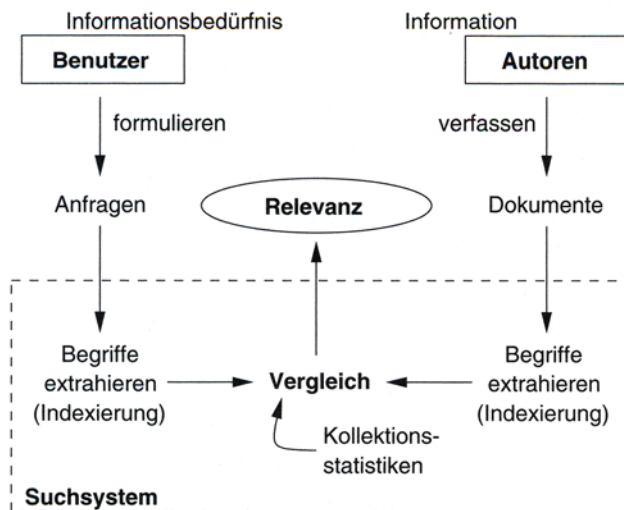
In der elektronischen Informationssuche wird die Beurteilung von Dokument-Relevanz automatisiert. Zwischen Autor und Nutzer steht kein beurteilender Mensch, sondern ein Suchsystem, das zu einer Anfrage die passenden Dokumente liefern muss.

Anstelle der Interpretation von Anfragen und Dokumente werden nur Begriffe im Text gefunden, gezählt und Zahlen miteinander verglichen. Anstelle des „Hintergrundwissens“ treten statistische Angaben über den gesamten Dokumentenbestand des Suchsystems sowie gewisse ‚Regeln‘. Resultat des Vergleichs ist ein Relevanzwert, der die Relevanz eines Dokuments gegenüber der Anfrage beschreibt → gefundene Dokumente werden anhand dieses Wertes in eine Liste mit fallender Relevanz sortiert.

Es wird deutlich, dass diese Automatisierung kaum adäquat eines menschlichen Experten funktionieren kann: Computer können die Suchanfragen nicht ‚verstehen‘, Statistik ist kein guter Ersatz für fehlende Interpretationsmöglichkeiten.

Semantisches Wissen fehlt Computern derzeit noch völlig, auch höhere Fähigkeiten zur Interpolierung und Korrektur von Suchanfragen. Umgekehrt kann kein Mensch wirklich ‚Computersprache‘ sprechen und Suchanfragen so formulieren, wie es ein Computer versteht.

Wichtig ist also, dass die Relevanzermittlung durch Suchmaschinen nur immer statistische Näherungen sind, die zudem noch in hohem Masse von den internen, nichttransparenten Regeln der SuMa abhängt und zunehmend durch bezahlte Treffer und ähnliche Wildwüchse überlagert wird.

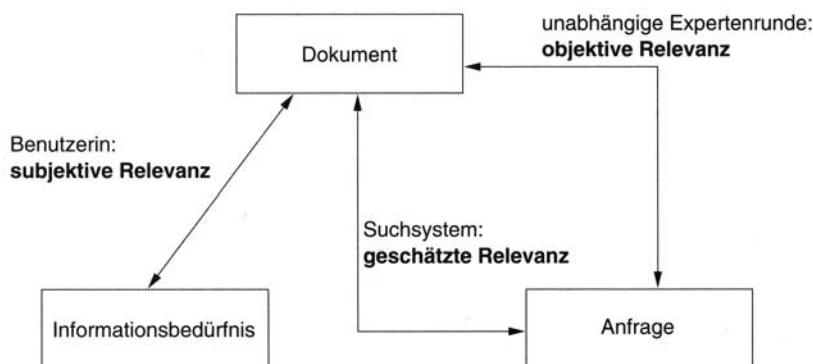


Als Problem kommt noch hinzu, dass Dokumente und Anfragen ja von Menschen mit persönlichen Schreibstilen und Wortwahlen verfasst sind. Dies beeinflusst aber die statistische Auswertung, die stark an den verwendeten Wörtern ausgerichtet ist. Die fehlende Fähigkeit, Synonyme semantisch zu erkennen, führt dazu, dass Suchsysteme somit zusätzlich grundsätzlich fehlerhaft sind. Das perfekte Suchsystem wird es somit niemals geben

1.2 Relevanz eines Dokumentes

Knapp gilt: Ein Dokument heißt relevant, wenn es bei einer Informationssuche gefunden werden soll.

Dabei gibt es unterschiedliche Formen von Relevanz:



Subjektive Relevanz: Ein Dokument ist für einen bestimmten Benutzer mit einem bestimmten, subjektiven Informationsbedürfnis relevant. Bzw. umgekehrt: Die Relevanz eines Dokumentes wird vom Benutzer aufgrund seines ganz persönlichen Interesses eingeschätzt. Dies wird aber von automatischen Suchalgorithmen kaum unterstützt werden, weil solche persönlichen Bedürfnisse kaum in wenige Stichworte formuliert werden können. *Beispiel: „Ich will für das nächste Referat eine möglichst genaue, aber allgemein verständlich formulierte Erläuterung der Ammoniaksynthese, am besten mit guten Abbildungen und etwa anderthalb DIN A4-Seiten lang.“*

Objektive Relevanz: Ein Dokument kann anhand von objektiven Faktoren / durch unabhängige Experten als relevant eingeschätzt werden. *Beispiel: Der Text zur Ammoniaksynthese kann beurteilt werden unter dem Gesichtspunkt Vollständigkeit des aktuellen Forschungsstandes, Güte der Literaturliste, Vollständigkeit der Darstellung (also z.B. Eignung für einen Enzyklopädiertext).*

Geschätzte Relevanz: Einem Suchsystem berechnet die Relevanz eines Dokumentes aufgrund von bestimmten, systeminternen Regeln. *Beispiel: Dokumente werden untersucht, wie oft der Begriff „Ammoniaksynthese“ im Dokument vorkommt, wie oft an prominenten Stellen, wie gut die Dokumente verlinkt sind ausgehend vom Wort „Ammoniaksynthese“ usw. usf.* Dies ist die einzige, automatisierbare Relevanzbestimmung.

Deutlich wird, dass die drei Arten der Relevanz kaum übereinstimmen.

Für den Nutzer am Wichtigsten ist der Fakt, dass die für ihn wichtige, subjektive Relevanz kaum gut mit der geschätzten Relevanz der SuMas übereinstimmen, aufgrund von:

- fehlende semantische Fähigkeiten (etwa bei mehrdeutigen Begriffen)
- Unterkomplexität der Suchanfragen
- Inkompatibilität zwischen menschlicher und Computersprache
-

1.3 Gewichtung von Dokumenten nach Relevanz

Wir haben also einige Probleme bei der Informationssuche kennen gelernt:

1. Häufig ist das Informationsbedürfnis aus der Anfrage eines Benutzers nicht klar ersichtlich.
2. Die Relevanz eines Dokuments bezüglich einer Anfrage ist schwierig zu beurteilen. Was das Suchsystem als relevant beurteilt, kann für den Benutzer subjektiv völlig irrelevant sein.
3. Fehlendes oder nicht aktiviertes Hintergrundwissen erschwert die Informationssuche zusätzlich.

Für dieses Problem gibt es keine Lösung sondern nur einen Workaround: Relevance Ranking.

Definition Relevance Ranking

Anordnen von Dokumenten gemäß absteigender Relevanz bezüglich einer Anfrage.

Ablauf Relevance Ranking

Erster Schritt: Nachdem der Benutzer eine Suchanfrage gestartet hat, werden alle verfügbaren Dokumente mit der Anfrage verglichen. Bei diesem Vergleich weist das System jedem Dokument einen Relevanzwert zu. Mit Hilfe des Relevanzwertes drückt das System die geschätzte Relevanz der jeweiligen Dokumente in Bezug auf die Anfrage aus. Je höher der Relevanzwert ausfällt, desto wahrscheinlicher stuft das Suchsystem ein Dokument bezüglich der Anfrage als relevant ein.

Zweiter Schritt: Nun sortiert das Suchsystem die Dokumente aufgrund des Relevanzwertes in absteigender Reihenfolge. Die so entstehende, geordnete Liste wird Rangliste genannt. Die Rangliste wird dem Benutzer präsentiert, der dann je nach ihrem Bedürfnis Dokumente auswählt und betrachtet. Durch die Sortierung der Dokumente in der Rangliste gemäß ihrer Relevanzwerte werden gleichzeitig völlig entgegengesetzten Bedürfnissen Rechnung getragen (bei nur einem Fakt reicht ein Treffer, bei einem umfassenden Überblick muss möglichst eine vollständige Dokumentmenge erstellt werden).

1.4 Rangierungsprinzipien

Um die Relevanz von Dokumenten bezüglich einer Anfrage zu berechnen, geht ein Suchsystem von folgender Grundannahme aus: Die Vorkommen von Suchbegriffen in einem Dokument geben Hinweise auf die Relevanz dieses Dokuments. Auf diese Annahme bauen statistische Modelle zur Berechnung auf, in der unterschiedliche Messwerte unterschiedlich stark gewichtet werden. Die Art der Messwerte und das ‚Mischungsverhältnis‘ sind von SuMa zu SuMa unterschiedlich und stellen eifersüchtig gehütete Geheimnisse der Betreiber dar.

Anhand von einigen besonders beliebten und effektiven Rangierungsprinzipien soll erklärt werden, welche Kriterien ein Suchsystem bei der Schätzung der Relevanz berücksichtigen kann. Weitere finden sich z.B. bei Karzauninkat 2005.

Rangierungsprinzip 1: Je mehr der Suchbegriffe in einem Dokument vorkommen, desto wahrscheinlicher ist das Dokument relevant.

Erläuterung: Dokumente, in denen alle Suchbegriffe auftauchen, werden höher gerankt als Dokumente, in denen nur einer der Suchbegriffe auftaucht.

Rangierungsprinzip 2: Je häufiger ein Suchbegriff in einem Dokument vorkommt, desto wahrscheinlicher ist das Dokument.

Erläuterung: 10 Vorkommen werden höher gerankt als 5.

Rangierungsprinzip 3: Dokumente, die seltene Suchbegriffe enthalten, sind mit einer höheren Wahrscheinlichkeit relevant als Dokument, die häufige Suchbegriffe enthalten.

Erläuterung: Fachbegriffe, Eigennamen, geographische Bezeichnungen etc. werden höher gerankt als übliche, statistisch häufige Wörter.

Rangierungsprinzip 4: Ein kurzes Dokument ist mit einer höheren Wahrscheinlichkeit relevant als ein langes Dokument, welches die gleichen Suchbegriffe gleich häufig erhält.

Rangierungsprinzip 5: Je näher die Suchbegriffe beieinander liegen, desto wahrscheinlicher ist das Dokument relevant.

Erläuterung: Bei der Suche nach „Big Ben“ wird ein Dokument mit der genauen Phrase „Big Ben“ höher gerankt als ein Dokument mit „Big [...n Worte...] Ben“

Rangierungsprinzip 6: Je früher die Suchbegriffe in einem Dokument vorkommen, desto höher ist seine Relevanz.

Erläuterung: Dokumente mit Suchworten zu Beginn werden höher gerankt als Dokumente mit Suchworten in der Mitte.

Rangierungsprinzip 7: Je öfter die Suchworte an prominenten Plätzen in einem Dokument vorkommen, desto höher ist seine Relevanz.

Erläuterungen: *Dokumente mit Suchworten* in Dokumenttiteln, Metatags, *Überschriften werden höher gerankt als Dokumente mit Suchworten im normalen Text.*

Diese Liste mit eher theoretischen Prinzipien ließe sich natürlich noch weiter fortsetzen.

1.5 Rangierungsprinzipien in der Praxis

Ein Suchsystem wählt die Dokumente zu einer Anfrage aufgrund der Rangierungsprinzipien aus. Das Resultat dieser Auswahl ist eine Rangliste, die an den Benutzer geschickt wird, aufgeteilt in kleine, darstellbare Häppchen. Auf der Hand liegt, dass die Rangliste nach abfallenden Relevanzwerten sortiert ist. Die vorgestellten Rangierungsprinzipien machen klar, nach welchen Kriterien ein Suchsystem die Dokumente theoretisch indizieren KANN. Allerdings sind die genauen Kriterien sowie die ‚Mischung‘ der Kriterien das Erfolgs- und darum das Firmengeheimnis der Suchdienste. Die genauen Prinzipien erfährt man zum Teil auf den Hilfeseiten der Suchdienste, zum Teil auf SuMa-Informationen-Seiten wie der „Suchfibel“, zum Teil entwickelt man als regelmäßiger Benutzer nach gewisser Zeit ein entsprechendes Gefühl.

1.6 Inhaltsunabhängige Bestimmung der Relevanz

Neben den bisher besprochenen Relevanzkriterien, die sich auf den Inhalt der Dokumente richten, gibt es Rangierungsprinzipien, die sich vom Inhalt lösen. Diese inhaltsunabhängigen Kriterien haben mittlerweile eine fast noch größere Bedeutung.

Rangierungsprinzip 8: Anzahl Referenzen. Für jedes Dokument wird gezählt, wie viele andere Dokumente darauf verweisen. Je mehr Referenzen ein Dokument auf sich vereinen kann, desto höher ist sein Relevanzwert. Dies ist ein recht erfolgreiches Kriterium, weil es die subjektive Relevanz vieler Nutzer statistisch auswertet und damit die subjektive Relevanz des Nutzers statistisch gesehen am nächsten kommt: Allgemein sehr beliebte (als relevant eingeschätzte) Webseiten sind vermutlich auch Webseiten, die beim Nutzer als relevant eingeschätzt werden. Dies geht selbstverständlich nur durch Kombination mit den inhaltsbasierten Kriterien. (Berühmt geworden durch Google, den berühmteste Missbrauch dieser Technik stellt konsequent auch die „Google-Bombe“ dar, siehe unten.) Diese Methode ist nicht von ungefähr erfolgreich, funktioniert die Bewertung von Fachbüchern / Fachartikeln / Fachautoren doch auf genau diesem Grundsatz. (Und wird von „Google Scholars“ für das Netz erfolgreich imitiert.)

Rangierungsprinzip 9: Anzahl der Zugriffe. Suchdienst kann recht leicht die Anzahl der Zugriffe auf ein Dokument ermitteln. Dokumente mit hohen Zugriffszahlen sind äquivalent zu Dokumenten mit hohen In-Links besonders relevant. Auch dieses Prinzip arbeitet nur kombiniert mit inhaltsbasierten Kriterien.

Rangierungsprinzip 10: Bezahlte Links. Zunehmend werden hohe Positionen in Ranglisten verkauft oder sogar versteigert. Mal recht deutlich markiert (wie bei

Google), mal sehr undeutlich verschleiert (und dadurch in höchstem Maße unseriös!) Damit wird die Relevanz eines Dokumentes anhand der Relevanz des Anbieters ermittelt und die subjektive Relevanz des Nutzers völlig ignoriert. Das Mittel ist als Refinanzierungsprinzip von Suchmaschinen an sich okay, muss aber deutlich markiert sein und ZUSÄTZLICH zu den nicht verkauften Ranglisten angeboten werden (gutes Beispiel ist auch Googles „AdWords“, das neben der eigentlichen Suche läuft).

1.7 Die dunkle Seite der Relevanzbestimmung: Webspamming

Suchen und Finden im Netz ist längst ein profitdominiertes Feld im Netz geworden: Wer als Firma unter den ersten zehn Treffern landet, wird angeklickt und verkauft, Treffer 11 bis 20 verkaufen nicht. Kein Wunder, dass die Manipulation des Indizierungsvorganges (unter dem euphemistischen Stichwort: Optimierung der Suchmaschinen-Position) ein gutes und intensives Geschäft ist. Dies aber führt zu einem Wettrennen zwischen Suchmaschinen-Betreiber, die zuverlässige Suchen anbieten wollen und Manipulationen der Spammer rigoros unterbinden und den Spammern selbst, deren Geschäft es eben ist, die Regeln der Suchmaschinen zu ermitteln und zu umgehen.

Längst funktionieren Manipulationen nicht mehr über einfaches, vielfaches Wiederholen von Stichwörtern oder das Anzeigen von beliebten Schlüsselbegriffen. (Solche Methoden werden mittlerweile konsequent bestraft durch schlechte Positionen.) Die derzeitigen Methoden laufen über z.B. Manipulationen der Verlinkung von Webseiten durch automatisch generierte Linkfarmen. (siehe weiter unten: Google Spamming).

Es wird deutlich, dass die Rankingmechanismen von Suchmaschinen mit höchster Vorsicht zu genießen sind. Keinesfalls sollten Ergebnisse EINER Suchmaschine kritiklos hingenommen werden!

2 Google und Relevanz

2.1 Google-Bombe

Unter einer Google-Bombe versteht man eine Manipulation des Google-Suchergebnisses für eine bestimmte Webseite durch vielfaches Setzen von Links mit einem vereinbarten Ankertext. Wenn nach diesem Text gesucht wird, zeigt Google nicht nur die Seiten mit dem Link an, sondern auch die Seite, auf die der Link verweist. Google-Bomben werden von Online-Communities eingesetzt, um Webseiten bestimmter Personen gezielt mit oft diffamierenden Schlagworten in Verbindung zu bringen.

Die erste Google-Bombe kam 2001 auf, als Adam Mathes die Worte talentless hack benutzte, um damit auf die Webseite eines Freundes zu verweisen. Er forderte andere auf, mit dem gleichen Text auf die gleiche Seite zu verweisen, und wenig später erschien die Webseite des Freundes bei Google auf Platz 1 -- unter den wenig schmeichelhaften Suchbegriffen talentless hack.

Die populäre Suchmaschine Google ist anfällig für diese Art der Manipulation, da nicht nur die Inhalte einer Webseite selbst für die bewertete Reihenfolge der Suchergebnisse entscheidend sind, sondern auch die darauf verweisenden Seiten, insbesondere die Verweise selbst mit ihren Ankertexten (siehe: PageRank).

Populär wurden die Google-Bomben 2003, als Gegner des amerikanischen Präsidenten George W. Bush sich absprachen und auf ihren Webseiten und Weblogs Links auf einen offiziellen biografischen Text über Bush setzten, jedes Mal

verbunden mit den Worten miserable failure (dt. erbärmliches Versagen, bzw. jämmerlicher Versager (z.B. in "He's a miserable failure.")). Die Google-Suche listete wenig später Bushs Seite als Top-Hit für eben diese Suchbegriffe.

Seitdem wurden Google-Bomben oft eingesetzt, unter anderem auch von Freunden des amerikanischen Präsidenten, die im Gegenzug versuchten, politische Gegner wie Jimmy Carter, Michael Moore und Hillary Clinton mit der Phrase miserable failure in Verbindung zu bringen. Inzwischen liegt mal Michael Moore, mal George W. Bush vorne. In Österreich bekannt wurde das Google-Bombing, als eine Suche nach völlige Inkompetenz zur heftig umstrittenen privaten Website des amtierenden Finanzministers Karl- Heinz Grasser führte.

Siehe auch: Suchmaschinen-Spamming, Linkfarm, wikipedia:Darwikinismus

Quelle: Wikipedia, <<http://de.wikipedia.org/wiki/Google-Bombe>>

2.2 Google AdWords

Adwords ist eine Produktlinie der Internet- Suchmaschine Google.

Mit diesem Namen werden bei Google buchbare Textanzeigen bezeichnet, die eine Ergänzung zum Suchergebnis darstellen sollen und bei Eingabe eines Suchwortes auf der rechten Ergebnisseite zu finden sind.

Im Gegensatz zu den üblichen Einträgen in Google ist die Einstellung von Textanzeigen nicht kostenlos. Der Inserent bezahlt hierbei einen Fixpreis für jeden Klick, der auf die Anzeige getätigt wird. Üblicherweise werden zu einem Stichwort bei Google maximal acht Textanzeigen angezeigt, die Position der Textanzeigen reguliert sich durch den Preis, den der Inserent bereit ist, für eine Textanzeige zu bezahlen. Je höher der gebotene Preis, desto besser die Position unter den Textanzeigen. Für jede Textanzeige muss der Inserent ein oder mehrere Stichwörter angeben, unter denen die Anzeige bei Google erscheinen soll.

Mit "Adwords" umschiffte Google die lange Zeit von Freunden der Suchmaschine propagierte Werbefreiheit. Bis dahin sollte sie zum einen ein Zeichen der Unabhängigkeit von Anzeigengeldern, zum anderen aber auch ein Garant für die schnelle Ladbarkeit der Seiten von Google sein. Die Adwords-Textanzeigen sind werbetaktisch nicht so aggressiv wie die Einblendung von Bannern, zudem präsentieren sie dem Benutzer der Suchmaschine mittlerweile oft bessere Resultate auf eine Suche, als es die Suchmaschine an sich kann.

Weblinks

- <http://adwords.google.de>
- <http://google-adwords.blogspot.com/>

Quelle: Wikipedia, <<http://de.wikipedia.org/wiki/Google%20AdWords>>

2.3 Google-Spamming

Mit Hilfe von Google kann man viele Besucher auf seine Webseite lotsen und mit entsprechender Werbung auch viel Geld verdienen. Google arbeitet ständig an der Verbesserung des Pagerank-Algorithmus um bekannte Schwachstellen zu beseitigen. Die Spammer sind Google dabei jedoch häufig einen Schritt voraus.

1. Google-Cloaking

Beim Cloaking existieren zwei Versionen einer Webseite, wobei an Google eine speziell optimierte Version geliefert wird. Google wird dabei anhand der IP-Adresse und dem HTTP USER AGENT erkannt. Links können auch mit NOFRAME oder mit CSS vor den Besuchern versteckt werden. Gleiche Schrift- und Hintergrundfarben werden als Spam erkannt.

2. Dynamische Webseiten und Parameterübergabe

Google wertet dynamische Webseiten mit Parametern (www.domain.de/?parameter=wert) ab. Deshalb nutzen Spammer das Apache-Modul Mode Rewrite, um die Parameter im Pathnamen (www.domain.de/parameter/wert.htm) zu verbergen. Weiterhin wird die Ausgabe der HTML-Seite gepuffert, um im Header die Größe in Bytes auszugeben und damit eine statische Seite zu simulieren.

3. Links mit hohem Pagerank erzeugen

Um in den Suchergebnissen unter die Top10 zu kommen, werden Links von Webseiten mit hohem Pagerank benötigt. Es gibt verschiedene Möglichkeiten, um Links von fremden Seiten mit hohem Pagerank zu erhalten. Diese können zum Teil automatisch per Skript erzeugt werden.

In Online Foren und Gästebüchern können Beiträge mit Spam-Links erstellt werden. Viele Webseiten erstellen automatisch Listen von Seiten, die einen Link auf sich gesetzt haben. Dies kann genutzt werden, um per Skript die Webseite mit einem HTTP REFERER auf die Spamseite aufzurufen. Den selben Effekt erzielt man, indem man die Referer-Seite als 1x1 Pixelbild am Seitenende einbindet.

4. Google-Circle

Um viele Besucher durch Google zu erhalten ist es gar nicht nötig, zu einem bestimmten Suchbegriff unter die Top10 zu gelangen. Eine breite Streuung an Suchbegriffen liefert auch viele Besucher. Um einen Google-Circle zu erhalten wird bei Google nach Suchbegriffen gesucht (z.B. Internet, Software, Reisen, Auto, etc). Die häufigsten 20 Wörter / Wortkombinationen aus den Suchergebnissen bilden dabei wiederum neue Suchbegriffe. Es kommt zu einer endlosen Kombination an Suchbegriffen und Trefferseiten.

Quelle: Wikipedia, <<http://de.wikipedia.org/wiki/Google-Spamming>>

Diskussion:

Was aber bedeuten die dargestellten Regeln und Mechanismen der Suchmaschinen für die Bewertung von Suchergebnissen und die Einschätzung der Qualität von Webseiten bei der realen Suche?

3 Sicheres Recherchieren

Die 8 goldenen Regeln der Suchmaschinenbenutzung

(nach: „Auf einen Blick: Die 12 goldenen Regeln kurz gefasst“, in: Machill 2005, S. 13-31, 41-43)

1. Verlassen Sie sich nicht auf eine Quelle

Verwenden Sie gerade bei komplexen Themenrecherchen immer mehrere Suchmaschinen. Denn: Unterschiedliche Suchmaschinen verwenden unterschiedliche Kriterien für ihre Rankinglisten. Somit können völlig verschiedene Websites als die „relevantesten“ aufgeführt werden. Vergleichen Sie die Auswahl, die Ihnen angeboten wird - und dann treffen Sie Ihre eigene.

2. Achten Sie auf versteckte Werbung

Werbung ist nicht immer als solche gekennzeichnet, sondern kann sich in den Ergebnislisten verstecken. Schauen Sie sich daher eine Ergebnisseite immer erst in Ruhe an, bevor Sie etwas klicken.

3. Hüten Sie sich vor den „Spammern“ im Netz

Das Wort „Spam“ wird bei den Suchmaschinen verwendet, wenn findige Webmaster ihre eigenen Websites so manipulieren, dass sie in den Ergebnislisten der Suchmaschinen höher eingestuft werden. Gehen Sie nicht denen auf den Leim, die durch diese Manipulation ihre eigenen vielleicht kommerziellen, politischen oder gar pornografischen „Botschaften“ unter die Leute bringen wollen. Deshalb: Formulieren Sie Ihre Suchanfragen möglichst genau und lesen Sie den Vorschautext gut durch. Sind dort nur Schlagwörter und keine Sätze zu sehen, verbirgt sich hinter der Webseite wahrscheinlich „Spam“.

4. Denken Sie bei der Auswahl aus der Suchmaschinen-Liste an Viren und den Schutz dagegen

Nicht jedes Suchergebnis ist vertrauenswürdig, nur weil es von einer Suchmaschine angezeigt wird. Manche Leute denken, dass sie schnell einen von der Suchmaschine angezeigten Link aktivieren können - und vergessen dabei die normalen Vorsichtsmaßnahmen, die im Internet sowieso gelten sollten. Das Internet ist auch ein globaler Tummelplatz, in dem sich wie auch im wahren Leben Betrüger oder schräge Spaßvögel finden.

Einen Virenschutz sollten Sie beim Surfen im Netz mindestens installiert haben, ebenso hilft die Installation einer Firewall.

5. Lesen Sie den Vorschautext gut durch

Oftmals mögeln sich unseriöse Angebote mit unerlaubten Praktiken auf die vordersten Ergebnisplätze in den Suchmaschinen. Der Vorschautext kann bei der Einordnung helfen: Er zeigt einen kleinen Textausschnitt der Seite, in dem die Suchbegriffe auftauchen. Lesen Sie diesen Text gut durch, denn oftmals erkennt man schlechte Ergebnisse bereits auf den ersten Blick.

6. Geben Sie Wucher-Websites keine Chance

Ein falscher Klick, ein falscher Tastendruck, und schon können 30 Euro und mehr fällig werden; es installieren sich „Dialer“ oder „Spyware“ auf Ihrem Rechner. Tippen Sie also nicht gutgläubig Antworten in Felder, die Sie nicht kennen, z. B.

„OK“. Dieses „OK“ kann nämlich Ihre Einverständniserklärung für einen „Mehrwertdienst“ sein.

7. Trauen Sie niemals der erstbesten Information - es könnten Verleumdung und Propaganda dahinterstecken

Manchmal verstecken sich rassistische Äußerungen dem Wolf im Schafspelz gleich in einem auf den ersten Blick seriösen Angebot. Prüfen Sie genau, wer für Informationen verantwortlich ist, bevor Sie sich Ihre Meinung bilden.

8. Nachrichtensuchmaschinen sind weder neutral noch vollständig

Noch viel stärker als in der normalen „Medienwelt“ sollte man sich bei den Nachrichtensuchmaschinen Gedanken über die Auswahl der Meldungen machen.

Denken Sie nicht, dass die Maschinen einen vollständigen Überblick über die Nachrichtenlage bieten. Längst nicht alle (Online-)Medien werden von den Suchrobotern der Nachrichtensuchmaschinen erfasst; außerdem reagieren die „News Search Engines“ stärker auf aktuelle Kurzmeldungen als auf Hintergrundberichte.