

Mario Fischer, Bastian Grimm

WHITE LINKHOUSE DOWN?

Backlinks sind für das Ranking bei Google wichtig. „Das wird auch so bleiben“, hört man allorts. Doch wie viel Zweckoptimismus steckt hinter solchen Aussagen? Wie viel davon ist kritik- und

DER AUTOR



Mario Fischer ist Herausgeber und Chefredakteur der Website Boosting und seit der ersten Stunde des Webs von Optimierungsmöglichkeiten fasziniert. Er berät namhafte Unternehmen aller Größen und Branchen und lehrt im neu gegründeten Studiengang E-Commerce an der Hochschule für angewandte Wissenschaften in Würzburg.

DER AUTOR



Bastian Grimm ist Partner bei der Berliner Fullservice Performance-Marketing Agentur Ads2people und verantwortet dort als VP of Search alle Aktivitäten & Kampagnen im Bereich SEO.

gedankenlos einfach der gängigen Meinung geschuldet und nach deren Mund geredet? Hilft gebetsmühlenartiges „Links sind nach wie vor wichtig“, auch Google langfristig davon zu überzeugen, dass man alles so lassen muss, wie es ist? Weil nicht sein darf, was näher besehen vielleicht längst überfällig ist? Das Ende der Backlinks als wichtiges Rankingsignal? Mario Fischer und Bastian Grimm haben sich darüber einmal losgelöst und abseits der gängigen Meinung strategisch Gedanken über eine mögliche Zukunft ohne Backlinkzählereien gemacht und versucht, den Gedanken-speer etwas weiter nach vorn zu werfen.

Für künstlich, schnell und billig aufgebaute Massenbacklinks (häufig auch als „Black Hat SEO“ bezeichnet), hat Google mittlerweile recht gut wirkende Gegenmittel im virtuellen Köcher. Aber was ist mit den sauberen Links, die von „White-Hat“-Suchmaschinenoptimierern ins Web gebracht werden? Setzt man einen Link, um den Besuchern einen Mehrwert zu bieten, damit sie eine oder die passende externe Webseite einfach per Mausklick finden können – dann ist das im Einklang mit den Richtlinien von Google. Wird der Link aber gesetzt, um das Ranking der angelenkten Seite positiv zu beeinflussen, stellt dies eine Verletzung dieser Richtlinien dar. Der Zweck ist also gut oder böse – nicht der Link. Aber woher und wie will Google diesen Zweck erkennen? Hierauf setzen diejenigen, die irgendwie weder weiß noch schwarz sind, sondern eher grau. Während es früher noch schwierig war, wirklich gut wirkende Links von guten und bekannten Websites zu bekommen, kann dies heute jeder mit genügend Geld einfach und schnell bewerkstelligen. Die Anzahl der Linkgeber wächst – nicht zuletzt durch die vielen Verlags- und Medienseiten, die gegen entsprechende Bezahlung zum Teil schamlos und ab und an sogar auch rechtswidrig ohne Anzeigenkennzeichnung Artikel mit gut getarnten Ausgangslinks mitten im Text platzieren. Wie lange kann das gut gehen?

Um zu verstehen, welches Szenario wahrscheinlich ist, muss man zunächst ein wenig ausholen. Nicht immer hilft der Blick in die Vergangenheit beim Erkennen der Zukunft – hier schon. Ebenso erhellend ist ein bewusst statischer Blick nach vorn.

Wenn sich nichts ändert – was wird passieren?

Angenommen, Google macht so weiter wie bisher und zählt fleißig Links zusammen. Wer viele gute Links hat und sonst auf den Seiten nichts Grundlegendes falsch macht, wird

auch weiterhin vorn ranken. Es ist bekannt, dass man sich schon heute praktisch beliebig und nur vom Geldbeutel beschränkt gute Links kaufen oder -immer noch- mieten kann. Diejenigen Websitebetreiber, die bisher auf freundliche Nachfrage hin kostenlos guten Content verlinkten, sind entweder zunehmend genervt von den überhandnehmenden Anfragen nach Linktausch oder -verkauf und lassen solche Anfragen ins Leere laufen, oder sie fragen sich, warum die Großen damit Geld verdienen dürfen und sie nicht, und verlinken fortan gegen Bezahlung. Es ist sicher nicht weltfremd, vorauszusehen, dass irgendwann ein nennenswerter Teil der Journalisten es ebenfalls satt haben könnte, für billiges Geld Content für große Plattformen zu erzeugen, und dies womöglich auf eigenen Domains zum Zwecke der Einnahmenerzielung über das Setzen von Links selbst in die Hand nimmt. Die Anzahl der Linkverleiher und -verkäufer wird ganz sicher ansteigen. Es wäre das erste Mal in der Geschichte der Menschheit, dass die meisten Individuen sagen: „Och nö, macht ihr anderen mal, ich muss da bei einem Boom mit leicht verdientem Geld nicht dabei sein.“ Edel sei der Mensch, Zwieback und gut.

Jetzt geht es erst so richtig los mit dem Linkverkauf

Wie viele in der Onlinebranche glauben auch heute noch, dass man in das Metatag „Keywords“ die Suchbegriffe für Google & Co. einträgt? Wie viele glauben, dass man einen Rankingvorteil von Google im unbezahlten Bereich bekommt, wenn man viel bezahlte AdWords schaltet? Wie viele

wissen noch immer nicht, dass Frames, Flash oder Ajax Probleme für Suchmaschinen darstellen? Wie viele verwenden die HTML-Überschrift-Tags H1 bis H5 zur Formatierung und nicht zur Auszeichnung von Gliederungsüberschriften und gerne auch noch immer in der falschen Reihenfolge? Wie viele Websites haben heute noch keinen Title und/oder kein Desc-Tag bzw. immer das Gleiche als Standardtext eingetragen? Wie viele Sitebetreiber werten die Seitenzugriffe noch immer mit Serverlogfiles aus, in denen die knapp 70 % maschineller Traffic und die eigenen Seitenaufrufe mit enthalten sind? Wie viele Webmaster haben noch immer den völlig nutzlosen, ja sogar kontraproduktiven Hinweis im Impressum, dass man für den Inhalt verlinkter Seiten nicht verantwortlich sei?

Die Liste ließe sich problemlos verlängern, genügt aber wohl, um zu zeigen, dass sich teilweise wichtige Informationen auch nicht innerhalb von Jahren bis zu den Verantwortlichen herumgesprochen haben. Wer ständig auf einschlägigen Konferenzen unterwegs ist oder Fachpublikationen liest, mag das kaum glauben. In der schnelllebigen Onlinewelt verbreiten sich News innerhalb eines harten Kerns oft in Minuten, aber bis sie wirklich die breite Masse oder gar Titel tragende Manager erreichten, vergeht offenbar doch sehr viel mehr Zeit, als man vermuten würde. Wenn man etwas tiefer mit den entsprechenden Tools recherchiert, kann man erkennen, dass sogar von Domains, die bereits von Google abgestraft wurden, noch immer fleißig Links gekauft und gesetzt werden. Noch nicht einmal dieses starke Signal hält Agenturen oder Sitebetreiber

Ist ein Link von einem bekannten Medienportal tatsächlich noch eine Empfehlung für eine Seite oder wurde er verkauft?

davon ab, weiterzumachen wie bisher. Ob Ignoranz, Unbekümmertheit oder schlicht Unwissen oder Unfähigkeit dahinterstecken, lässt sich schwer beurteilen. Besonders schlau wirkt es jedenfalls nicht.

Man darf also getrost davon ausgehen, dass Linkkauf immer gefährlicher wird und immer öfter zu Strafen führt, sich diese Tatsache aber auch nächstes und übernächstes Jahr noch nicht breit rumgesprochen haben wird. Man muss sich ja nur ansehen, wie viele Angebote bzw. Agenturen aus dem Boden gestampft werden, die genau dies anbieten – den Backlinkaufbau zu übernehmen. 50 Links in zwei, fünf oder zehn Tagen? Kein Problem. Goldgräberstimmung – und es werden eben ja erst die Schaufeln verkauft. Dass der Hauptteil des Goldes schon lange weggeräumt sein könnte, wird verdrängt oder ignoriert. An der Börse verkaufen die Profis, wenn die Hausfrauen (was keine Herabwürdigung dieser Tätigkeit sein soll) anfangen, Aktien zu kaufen. Erleben wir so eine Blase gerade beim Linkverkaufen?

Wer auf Eiern läuft, sollte besser nicht hüpfen

Nun, nehmen wir – wie oben beschrieben – einmal an, immer mehr Unternehmen kauften Links zu oder ließen sie von Agenturen aus dubiosen Quellen erstellen. Google wäre gezwungen, immer mehr Domains abzustrafen und von den vorderen Plätzen zu verbannen. Damit würde es aber zunehmend auch völlig seriöse Websites erwischen, deren Besitzer tatsächlich nicht wissen, auf welch dünnes Eis sie durch Agenturen geschoben werden oder bereits wurden.

Es würde immer mehr wertvoller Content verschwinden bzw. wäre nicht mehr zu finden, noch besser, nicht mehr zu „ergoogeln“ sein. Das Web könnte inhaltlich deutlich ärmer werden. Zarte Ausläufer sehen wir

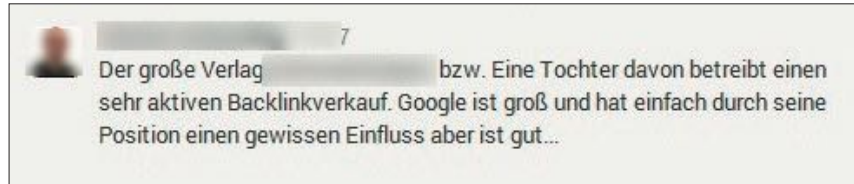


Abb. 1: Im Web wird fleißig diskutiert, welcher Verlag Links verkauft – teilweise recht aggressiv



Abb. 2: Am 14. März verkündete Matt Cutts, dass ein deutsches Linknetzwerk bestraft würde

schon heute. So erhielten z. B. große Reise- oder Gesundheitsportale durch Linksünden in der Vergangenheit empfindliche Strafen und ihre Seiten bzw. die angebotenen, durchaus oft guten Informationen sind praktisch nicht mehr über Index auffindbar. Wenn nun immer mehr Links von immer mehr Sitebetreibern gekauft werden, müsste es auch immer mehr solche Strafen geben. Immer mehr gute Seiten würden verschwinden.

Wäre diese Situation, dass die Suchergebnisse aufgrund von Strafaktionen schlechter werden, aus Googlesicht tolerierbar? Früher griffen nur Spammer zu solchen Maßnahmen (künstlicher Linkaufbau), deren minderwertigen Inhalt niemand vermisste. Wenn nun auch seriöse Unternehmen die gleichen Mittel verwenden, dann fällt eben auch Hochwertiges weg, und das kann niemand wollen. Am wenigsten Google selbst.

Soll Google als Ausweg dann einfach die Strafschwellen für künstliche Links lockern? Das wäre sehr unwahrscheinlich. Als einziger Ausweg bliebe wohl, den Linksignalen weniger Gewicht zu geben oder sie ganz am

Ende fast gänzlich zu ignorieren.

In der letzten Ausgabe hatten wir berichtet, dass Matt Cutts von Google am 14. März per Twitter bekannt gab, man habe ein deutsches Linknetzwerk entdeckt und werde dagegen vorgehen. Zwischen dem 16. und 17. März wurde der sichtbare PageRank* einiger Zeitschriftenportale deutlich nach unten gestuft (für einige Beispiele siehe Abbildung 3). Die ungewöhnliche Stärke der Abstufung (teilweise von PageRank 7 auf 3) sowie die auffällige zeitliche Nähe könnte einen Zusammenhang vermuten lassen, der dadurch natürlich aber weder zwingend noch für die hier beispielhaft herausgegriffenen Domains bewiesen ist. Wie in der Balkengrafik zu sehen ist, wird für die links der Trennlinie ausgewiesenen Domains mittlerweile (Stand 14.05.2014) wieder der alte PageRank-Wert ausgewiesen. Einige der anderen Domains (rechts) stehen allerdings noch immer auf dem abgestuften Niveau.

Man kann jetzt spekulieren, ob einige dieser Domains schnell und professionell reagierten und der Grund für die Beanstandung behoben

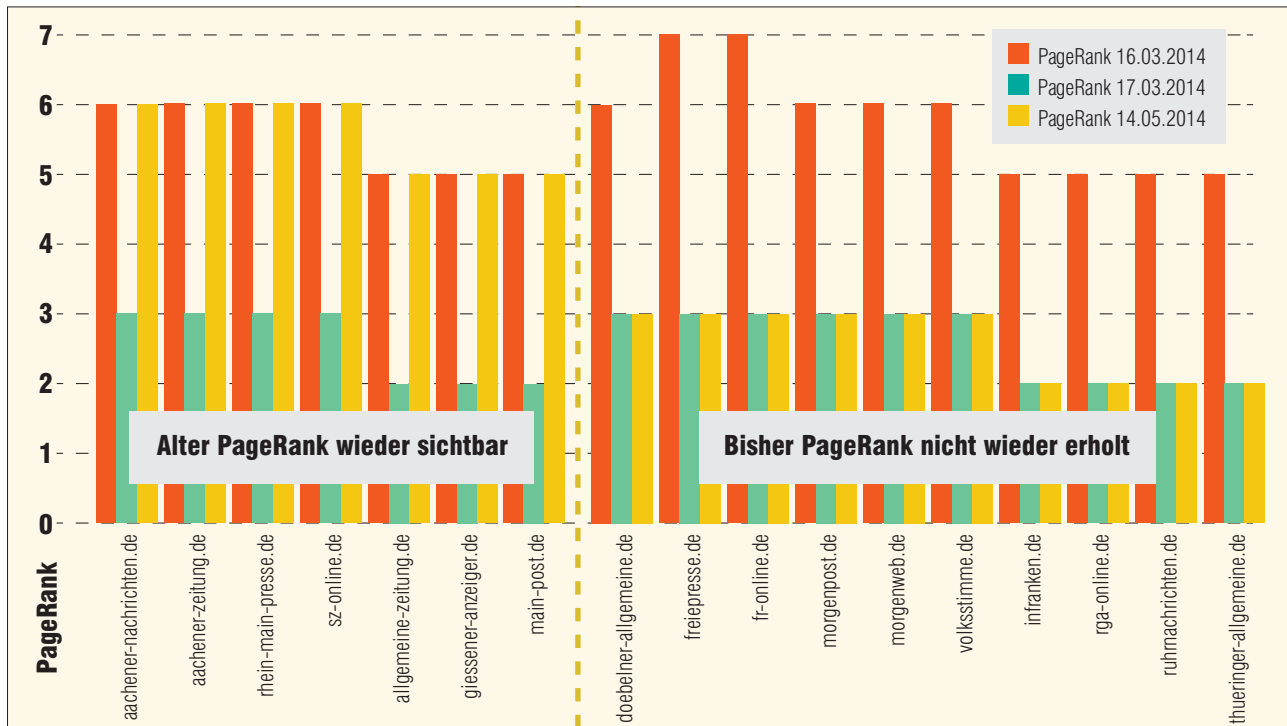


Abb. 3: Vom 16. auf den 17. März wurde offenbar bei einigen deutschen Nachrichtensites der PageRank massiv zurückgestuft

wurde und ob die anderen dies bisher unterließen. Allen mittels PageRank abgestraften Domains ist allerdings eines gemeinsam: Das eigene Ranking veränderte sich nach der Herabsetzung nicht nennenswert. Sieht man sich als Beispiel fr-online.de an, wo die Domain durchaus drastisch von PageRank 7 auf den Wert 3 herabgestuft wurde, und vergleicht die Rankings über die entsprechende Zeit sicherheitshalber in mehreren SEO-Tools (Abbildung 4), ergibt sich für diese spezielle Woche – nichts. Im Gegenteil, hier stieg die Sichtbarkeit kurz danach sogar leicht. Dies lässt vermuten, dass die Verminderung des PageRanks wohl generell eher als Signal an die Betreiber zu interpretieren ist und vielleicht auch als Signal an potenzielle Linkkäufer, die ja noch immer stark auf diese Zahl schielen, wenn es um den zu zahlenden Gegenwert in Geld für einen solchen Link geht. Viele Unterseiten einer PageRank-7-Domain werden oft noch mit dem Wert 5 oder 4 in der Toolbar ausgewiesen. Dafür kann man deutlich (!) mehr Geld verlangen, als wenn dort eine 1 oder 2 stünde, die praktisch auch jede Feld- und Wiesenseite ausweist.

Vielleicht ist es aber auch politischen Gründen geschuldet, dass ausgerechnet bei Sites von Verlagen bei Verletzungen der Richtlinien von Google nicht wie sonst üblich eine Rückstufung aus dem Index erfolgt? Auch darüber wurde in der letzten Ausgabe von Website Boosting ja spekuliert: welche publizistische Reaktion Google von Zeitungen und Zeitschriften zu erwarten hätte, wenn man das dort oft unglaublich hemmungslose Linkverkaufen bestrafen würde. In der Öffentlichkeit würde der wahre Grund dafür wohl weder kommuniziert noch verstanden. Viel leichter glaubhaft zu machen wäre, dass Google sich damit ganz profan an Verlagen wegen deren ständiger Nadelstiche gegen den Suchkonzern rächen würde. Und genau das hätte man der breiten Masse wohl auch so verkauft.

Ein Grund mehr für die Einsicht, wie viel Ärger das Thema Backlinks und der daraus resultierende Verkauf selbiger Google mittlerweile einbringt. Nur am Rand sei noch Folgendes erwähnt: Der Redaktion liegen mittlerweile Nachweise dafür vor, dass Tochterunternehmen von Verlagen sogar

per Kaltakquise via Telefon bei Web-sitebetreibern anfragen. Das Problem wurde also durch die öffentlichkeitswirksamen Strafmaßnahmen offenbar in gar keiner Weise eingedämmt, im Gegenteil, es weitet sich sogar noch aus. Und in Teilen erscheint das aktuell sogar noch logisch: So lange man doch nur den PageRank verliert, aber ansonsten keinen Traffic, wird eben munter weiter verkauft.

Zu wenig externe Signale für „Longtail*“-Suchen

Laut Google sind etwa 20 % aller Sucheingaben neu, wurden also noch nie so eingegeben. Bei 40.000 Suchanfragen weltweit pro Sekunde kann man sich eine Vorstellung machen, wie differenziert Suchanfragen gestellt und wie viele unterschiedliche Worte in nahezu unendlichen Kombinationen verwendet werden. Das bedeutet, dass für rund jede fünfte Suchanfrage bei Google keine historische Erfahrung z. B. in Form von Klickraten, Bounce-Rates oder Bewegungsdaten auf dem Ziel (über den Chromebrowser messbar) vorliegen, auf die man zur Qualitätsprognose und -sicherung

*siehe Online-Glossar unter www.websiteboosting.com



Abb. 4: Die Sichtbarkeit blieb trotz PageRank-Strafe erhalten (Tools v. l. n. r. Sistrix, Xovi, Seachmetrics, SEOlytics)

zurückgreifen könnte. Dazu kommt der Umstand, dass ein großer Teil aller Backlinks –sofern sie natürlich entstanden – in der Regel immer auf die Startseite einer Domain gehen und nicht direkt auf eine Unterseite. Es liegen also meist auch keine verwertbaren (externen) Linksignale für die Unterseiten vor. Es bleibt nichts anderes übrig, als dem Content der Seiten das Hauptgewicht zu geben. Früher waren das einfachere Signale wie die Keyworddichte im weitesten Sinne, Fettschrift, Aufzählungen, Wortpositionen und einige andere mehr. Heute ist Google sehr viel weiter und kann tatsächlich mehr und mehr den semantischen Zusammenhang verstehen. Selbstverständlich muss auch eine Rolle spielen, auf welcher Domain diese Inhalte erschienen. Dieser Domain-„Trust“ wird nach Meinung von Experten immer wichtiger.

Zwischenfazit:
OnPage-Signale waren bei Longtail-Suchen schon immer wichtig und könnten zukünftig noch wichtiger werden.

Dass Google denjenigen Websites weniger traut, die viele Links

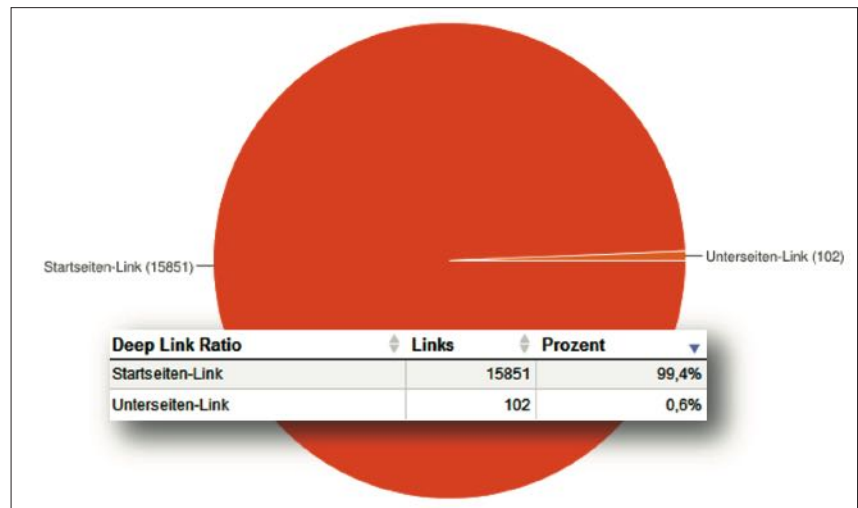


Abb. 5: Ein nicht manipuliertes Backlinkprofil sieht nicht selten so aus – wenig Links gehen auf Unterseiten (Quelle: LinkResearchTools)

auf Unterseiten haben, bei denen die Ankertexte optimierte und werthaltige Keywords enthalten, wurde vielen erst durch das **Penguin-Update*** klar. Das wurde u. a. ja genau dafür ausgerollt, um Seiten mit übertriebenen SEO-Linkmaßnahmen in der Rankings weniger Chancen zu geben.

Google schreit uns förmlich an, gute Inhalte besser zu strukturieren

Schema.org, Microdata, RDFa – ja sogar einen „Data Highlighter“ stellt man nicht technikaffinen Webmastern

zur Verfügung, damit diese ihren Websites per einfachstem Drag and Drop eine Datenbeschreibung und damit auch eine Struktur verleihen können. Durch das Hummingbird-Update im letzten Jahr und nicht zuletzt über den Knowledge Graph kann man unschwer erkennen, wie weit Google im Deuten von Zusammenhängen schon ist. Die Maschine weiß, dass Westerstiefel in der Regel aus Leder sind, etwas mit Mode und/oder Reiten zu tun haben, von Menschen gekauft und getragen werden, etwas mit der Entität „Fuß“ zu tun hat, häufig verschiedenen Zah-

*siehe Online-Glossar unter www.websiteboosting.com

Organization of Schemas

The schemas are a set of 'types', each associated with a set of properties. The types are arranged in a hierarchy.

Browse the full hierarchy:

- One page per type
- Full list of types, shown on one page

Or you can jump directly to a commonly used type:

- Creative works: [CreativeWork](#), [Book](#), [Movie](#), [MusicRecording](#), [Recipe](#), [TVSeries](#) ...
- Embedded non-text objects: [AudioObject](#), [ImageObject](#), [VideoObject](#)
- Event
- Health and medical types: notes on the health and medical types under [MedicalEntity](#).
- Organization
- Person
- Place, [LocalBusiness](#), [Restaurant](#) ...
- Product, [Offer](#), [AggregateOffer](#)
- Review, [AggregateRating](#)
- Action

We also have a small set of primitive data types for numbers, text, etc. More details about the data model, etc. are available [here](#).

Abb. 6: Schema.org hält bereits heute viele Typen von Metadaten bereit

len zugeordnet sind (die Größen) und Worte auftauchen, die eine Farbe darstellen. Man weiß, dass sie mit Läden zu tun haben, in denen sie verkauft werden. Häufig enthalten Webseiten auch ein Währungssymbol, vor oder nach einer Zahl mit zwei Kommastellen. Worte wie „Ziernaht“, „Absatz“, „rahmengenäht“, „Rindsleder“, „robust“, „bequem“ und einige andere mehr tauchen sehr häufig als sog. Kookkurrenz (vulgo: gemeinsam) auf, wenn es um Westernstiefel geht.

Mittels solcher Metainformationen ist es Google möglich, Seiten, bei denen es wirklich thematisch um Westernstiefel geht, von solchen zu unterscheiden, die nur das Wort mehrfach verwenden. Zum Teil hilft dabei das Auszeichnen mit den oben erwähnten Metadaten, zum großen Teil werden solche Zusammenhänge aber z. B. durch Tripelstores (damit werden Beziehungen zwischen Entitäten gespeichert) selbst ermittelt.

Google weiß, ob und wie gut wir tatsächlich sind

An welchem Signal könnte man wohl gut erkennen, ob eine Domain von den Nutzern gemocht wird oder nicht? Wahrscheinlich ist die Besuchshäufigkeit und auch -dauer ein guter Indikator, und auch der Zugangsweg. Wird der Domainname direkt in den Browser eingetippt, dann ist die

Domain bekannt. Wird der Domainname bzw. der Brand in den Suchschlitz eingegeben und dann auf das Ergebnis mit der Domain geklickt, gilt das ebenso. Kommen Besucher über einen Link auf einer anderen Site dorthin, dann ist dieser Link wohl nützlich. Oder wird eine Domain hauptsächlich durch die Keywordsuche bei Google gefunden und es fehlen solche vertrauenswürdigen Elemente wie oben? Aus dem Verhältnis dieser (und anderer) Zugangswege kann man statistisch recht gut ableiten, ob und wie viel Vertrauen man einer Domain algorithmisch zuweisen kann. Für Google ist es sogar noch einfacher. Sie müssen einfach nur ein Set sehr guter und vertrauenswürdiger Domains bestimmen und

herausmessen, wie dort die Zugangsverhältnisse sind. Dies entspricht dann keiner Vermutung, sondern der Realität der Bewegungen der Surfer -und signifikante Abweichungen selbiger lassen sich relativ einfach erkennen.

Woher hat Google diese Daten? Es werden immer wieder Vermutungen laut, man nähme dort einfach die Daten von Google Analytics. Unabhängig davon, dass dies von Google immer bestritten wurde, macht es auch wenig Sinn. Es nutzen weit mehr Menschen den hauseigenen Chromebrowser und über diese Bewegungsdaten (deren Übermittlung sich übrigens auch abschalten lässt) lassen sich Bewegungen durch das Web sehr viel besser hochrechnen. Hier stehen komplette Usersessions über alle Webseiten zur Verfügung, Analytics zeigt nur einen Ausschnitt, und wenn jemand auf eine Domain ohne Analyticscode ginge, wären diese Sessions mit unbekanntem Ziel unterbrochen.

Und das Gegenteil von gut? Auch das ist leicht über diesen Weg herausmessbar. Die Domains oder Seiten mit (zu) kurzer Besuchsdauer und vor allem die, bei denen überproportional viele Suchende sofort den Backbutton drücken und ein anderes Ziel aus den weiteren Suchergebnissen auswählen

„As the name implies, Chrome Data Compression is a Chrome-only feature available for Android and iOS platforms. When enabled, unless the request is made from an Incognito window or is going to an HTTPS site, it is routed to the compression proxy which then fetches the content, optimizes it, and returns the content to users device“

Ilya Grigorik, Google (Quelle: <http://einfach.st/igri>)

(sog. Bounces), die interessieren wohl niemanden wirklich – zumindest für ein bestimmtes Suchwort, vielleicht aber auch überhaupt! Denn Seiten, die in Summe von vielen Besuchern jeweils nur kurz im Browser geöffnet bleiben, deren Inhalt mag wohl niemand so besonders. Dazu kommen noch viele weitere Möglichkeiten, Bewegungen auf Webseiten auch über andere Browser zu messen, z. B. über Cookies von AdWords, Remarketing, AdSense, DoubleClick, Publisher-Tags, G+-Buttons, Google Maps oder auch einfach nur über Schriftarten, die beim Aufruf einer Seite von Google-Servern zur Verfügung gestellt werden. Das neueste Projekt „Data Compression Proxy“ macht den Webseitenabruf auf Android-Geräten schneller – aber damit auch alle mobilen Bewegungen erfass- und messbar. Schweben demnächst noch Ballons mit Internetzugang für alle im Himmel („Google Loon“) oder denkt man an das Google-Fiber-Projekt, dann wird klar, dass solche Bewegungsdaten praktisch an allen Ecken und Enden immer umfassender und detaillierter anfallen und ausgewertet werden können.

Die Sprünge von Seite zu Seite, von Domain zu Domain sind also für Google zumindest statistisch hochgerechnet relativ transparent.

Zwischenfazit:
Die Beliebtheit einer Domain
oder die Werthaltigkeit des
Contents einer einzelnen Webseite
und warum bzw. über welche
Suchbegriffe die Besucher
dorthin gehen, ist Google
bestens bekannt.

Links werden zunehmend automatisiert gesetzt

Wie viele Links werden domain-übergreifend eigentlich noch bewusst manuell gesetzt und wie viele automatisch von einem System (CMS, Blog-

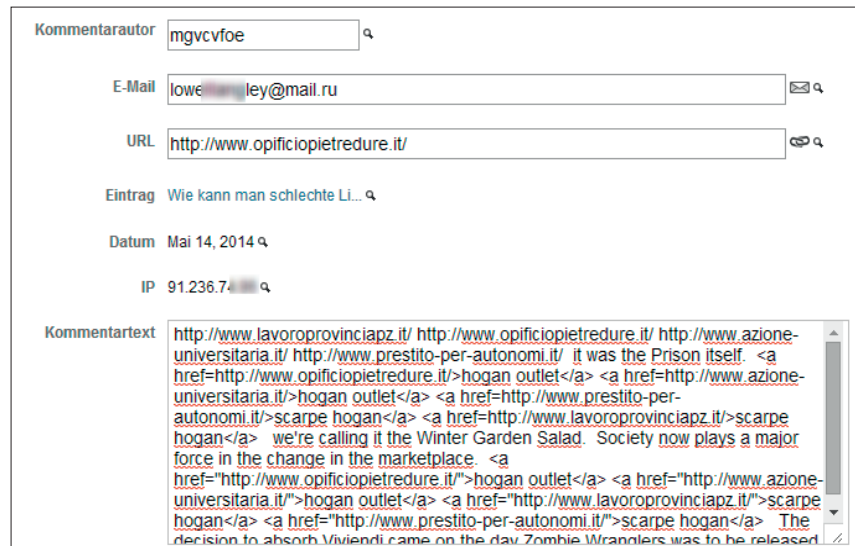


Abb. 7: Typischer Kommentarspam, wie er teilweise mehrmals täglich automatisiert auf deutschen Blogs eingeht

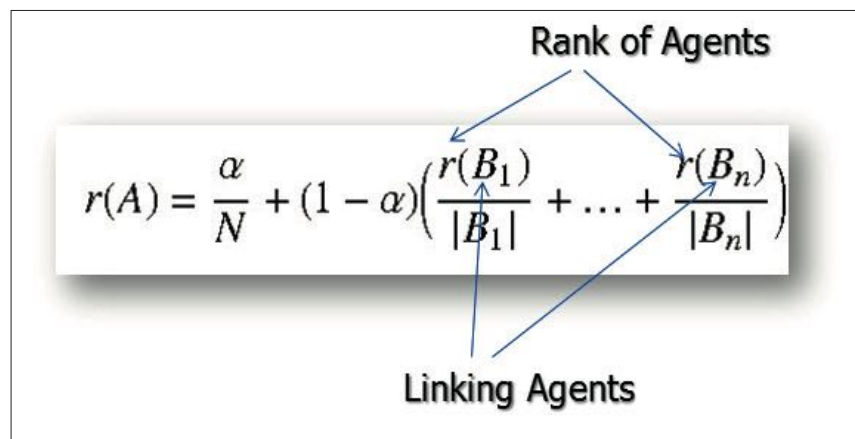


Abb. 8: Die zentrale Formel für die Kalkulation eines „Reputation Scores“ bzw. der Linkkalkulation über „trusted agents“ aus dem US-Patent US2011/0289095 A1 von Google, das 2011 genehmigt wurde

software, Pingbacks etc.)? Gerade auch in sozialen Netzwerken ist nicht immer Contentqualität der Grund für solche Links – man denke nur beispielhaft an die überproportional häufig gemochten Katzenbildchen. Dazu kommt noch ein massives „Oversharing“, weil Postings übernommen werden und zugleich auch nicht selten automatisiert in mehreren Netzwerken auftauchen.

Und kann man diese automatischen Links von bewussten maschinell trennen? Oft stellen solche Links ja auch nur eine Art „Momentanempfehlung“ dar, bleiben aber praktisch auf ewig im Netz.

Rel=“author“ und rel=“publisher“: Dir kann ich vertrauen?

Für Google tut sich zunehmend auch eine andere Quelle zur Ermittlung von Vertrauen auf: das Google+-Netzwerk. Dort agieren Menschen („author“) und Unternehmen („publisher“) und durch den sog. Social Graph ist man sehr gut in der Lage, echte Accounts von Spammer-Accounts zu unterscheiden. Die Verwurzelung eines normalen Users über Beziehungen und Interaktionen zu anderen Usern ist in der Regel sehr viel tiefer und intensiver und damit messbar. Dazu kommt, dass

„**Ein wirkliches Expertenprofil bei Google+ aufzubauen, dürfte wohl derzeit das am aufwendigsten zu erreichende und am schwierigsten zu fakende Signal sein.**“

man algorithmisch recht gut auch fachliche bzw. themenbezogene Expertisen herausrechnen kann. Der Fotoexperte ist von einem Liebhaber von Oldtimern oder einem Uhrenspezialisten unterscheidbar.

Entsprechende Patente von Google legen diese Vermutung durchaus nahe. Verbindet sich nun ein solcher Uhrenspezialist über das rel="author"-Tag mit einer Webseite, geht damit ein eindeutiges Signal an Google. Wenn genügend solcher Signale vorliegen und die Qualität sichergestellt ist – was liegt näher, als einer solchen Webseite entsprechend mehr Trust für ein spezielles Thema zuzuweisen? Oder anders betrachtet: Warum sollte Google auf solche sehr schwer zu fakende Signale verzichten? Backlinks kaufen geht vergleichsweise einfach. Aber ein wirkliches Expertenprofil aufzuziehen, dessen „Gültigkeit“ von vielen anderen Usern über den Social Graph und entsprechende Interaktionen (häufige, längere und thematisch passende Kommentare, Erwähnungen, Circles, Plus-button etc.) ständig bestätigt wird, das ist ungleich mehr Aufwand, braucht sehr viel Zeit und vor allem eines: wirkliche fachliche Erfahrung!

Werden Links als Signal künftig wegfallen?

So weit wird es sicherlich auf absehbare Zeit nicht kommen. Aber man kann wohl mit hoher Wahrscheinlichkeit davon ausgehen, dass sich die derzeit noch dominante Wirkung von Backlinks stark abschwächt. Zudem

wird wohl die Zahl der Links, die tatsächlich überhaupt noch gewertet werden, weiter stark zurückgehen.

Bereits jetzt wird womöglich ein großer Teil an Backlinks ignoriert oder sogar (nur) zur Ermittlung der Spamwahrscheinlichkeit verwendet. Vorstellbar und gar nicht so unvernünftig wäre es, z. B. nur die Links zur Bewertung zu nutzen, die tatsächlich von einer bestimmten Anzahl an Usern durch Anklicken verwendet werden. Die Besucher aktivieren quasi einen Link, der dann erst auf die Link empfangende Seite ein positives Signal erzeugt. Links von Seiten ohne Besucher ganz einfach zu ignorieren, ist vergleichsweise leicht für Google und wird wohl mit ziemlicher Sicherheit bereits heute schon praktiziert.

Die tatsächliche Linknutzung würde damit die im ursprünglichen PageRank-Modell abgebildete theoretische Nutzungswahrscheinlichkeit des „Random Surfer“ ersetzen. Damals konnte man diese Daten nur durch Linkbeziehungen errechnen, heute liegen, wie oben ausgeführt wurde, die tatsächlichen Werte vor.

Dieses durchaus plausibel wirkende Gedankenspiel lässt sich noch weiter treiben: Warum dann nicht auch einen Seitenbetreiber von der Reputation her gesehen abwerten, wenn er viele Links anbietet, die von niemandem geklickt werden? Vorausgesetzt, eine Seite hat genügend Traffic, ein Großteil der dortigen Links wird aber nicht genutzt – dann könnte man durchaus argumentieren, dass er wohl Dinge anbietet (hier externe Links), die von Besuchern als unnütz erachtet werden.

Bereits heute empfehlen Experten Sitebetreibern, die aufgrund falscher Verlinkung eine manuelle Strafe bekommen, neben minderwertigen Links gerade auch solche entfernen zu lassen oder in das Disavow-Tool einzupflegen, über die kein Traffic auf die eigene Website kommt. Im Webanalytics-

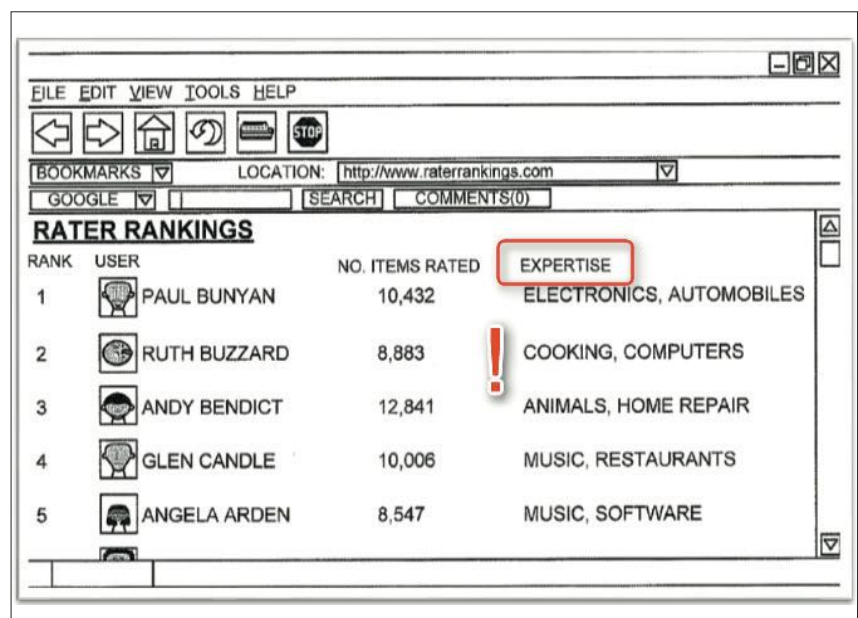


Abb. 9: This „... is a diagram of an exemplary graphical user interface ... that may provide user ranking information“ (US-Patent US 2011/0041075 A1, Google Inc.)

Tool kann man ja genau ersehen, von welchen Quellen tatsächlich Besucher kommen. Filtert man Direkteingaben der Domain und Traffic von Suchmaschinen weg, bleiben die Domains stehen, von denen Besucher über Links auf die eigene Site kommen. Die Königsklasse ist, nun noch diejenigen auszufiltern, welche die Site sofort wieder verlassen haben (Bounce). Wie bereits oben erwähnt, kennt Google diese Werte – es macht daher keinen Sinn, den Kopf in den Sand zu stecken und dieses (negative) Besuchervotum zu ignorieren und weiterhin tumb die bloße Anzahl an Visits als Erfolgsmaßstab zu zählen.

Wie werden die Spammer darauf reagieren?

Die Frage der Reaktion von Betroffenen bei Regeländerungen wird allzu oft ignoriert. Das Einbauen von Wegfahrsperrern hat die Aktivitäten der Automafia nicht lange davon abgehalten, sich wertvoller Fahrzeuge zu bemächtigen. Sollte das bloße Linksetzen als vergleichsweise einfach zu erzeugendes Signal zunehmend entwertet werden, wird der schwarze oder zumindest dunkelgraue Teil der Branche wohl darauf reagieren. Denkbar wäre zum Beispiel, Benutzern Programme unterzujubeln, die unsichtbar im Hintergrund Webseiten öffnen oder den Klick auf Links simulieren, um breit verteilt entsprechende Signale zu erzeugen. Traffic derart zu fälschen, ist allerdings lange nicht mehr so trivial wie z. B. automatisiert Content zu scrapen (von andern Seiten zu kopieren), zu vermischen und versehen mit Links in das Web zu stellen. Aber mit solchen trojanischen Programmen wird man wohl zukünftig rechnen müssen. Ein kleiner Teil der wirklich bösen Black Hats arbeitet auch heute schon mit ähnlichen Methoden. Dabei geht es nicht nur um die Manipulation von Suchergebnissen, sondern vor allem

INFO

Dass die Bewertung von Links seitens Google irgendwann komplett außer Kraft gesetzt wird, ist eher unwahrscheinlich. Dem Vernehmen nach experimentiert man in Mountain View aber immer wieder, wie Suchergebnisse ohne Linkwertung aussehen würden. Die so simulierten Ergebnislisten sind aber wohl noch spürbar schlechter, als die aktuell sichtbaren. Bei der semantischen Interpretation von Seiteninhalten und deren Zuordnung zur wirklichen Suchabsicht, steht man allerdings eben auch erst noch am Anfang. Insofern könnte es tatsächlich nur eine Frage der Zeit sein, bis Google linguistisches Wissen in technische Algorithmen umgegossen hat, um das Gewicht von insbesondere manipulierten Linksignalen mehr und mehr abschwächen zu können. Dies wird sicherlich für Deutschland nicht ohne Vorwarnung kommen, denn aufgrund der einfacher erkennbaren Sprachmuster wird man damit wohl zuerst in den USA arbeiten. Dass Google bei der Inhaltserkennung englischer Sprache schon sehr viel weiter ist, erkennt man bereits jetzt schon unschwer bei Vergleichen beim sog. Knowledge Graph. „Deutsche Sprache – schwere Sprache“ – das dürfte uns hierzulande sehr wahrscheinlich einen zeitlichen Vorlauf, ähnlich wie beim [Panda Update*](#), bescheren.

auch um das Fälschen von Metriken per Forced Click, manipulierte Toolbars und Browser-Plug-ins, das Austauschen von Affiliate-Links und/oder -Cookies und dergleichen mehr. Wer sich näher mit dem Thema „Black Hat CRO“ beschäftigt hat oder beschäftigen muss, weiß ein Lied davon zu singen, wie kreativ Betrüger sein können, wenn es um die Vorspiegelung falscher Tatsachen im Web geht.

Allzu fleißige Suchmaschinenoptimierer haben beim Blick auf die Vergangenheit viele von Google verwendete Signale so überstrapaziert, dass sie heute praktisch nichts mehr zählen. Die generelle Reputation von Blogs, insbesondere die Kommentarfunktion, wurde ruiniert, die Anzahl an Keywords auf Ankertexten, die Anzahl an Deeplinks, Bookmarkdienste, Webkataloge, Infografiken, die Geschwindigkeit des Linkzuwachses und vieles mehr. Warum sollten sie nicht auch längerfristig die Wirkung von Backlinks ruinieren?

Da bei Google sehr fähige Denker sitzen, darf man getrost davon ausgehen, dass man diese Entwicklung bereits seit Langem voraussah und an Gegenmaßnahmen arbeitet. Bis die treffsicher sind, vergeht wegen der Komplexität oft vergleichsweise viel

Zeit. Von einem kann man aber mit Sicherheit ausgehen: Lange wird man der Manipulation per Backlinks wohl nicht mehr tatenlos und ohne Gegenmaßnahmen zusehen. Eine weitere Abwertung der Wirkung von Backlinks und am Ende vielleicht sogar irgendwann das völlige Ignorieren derselben erscheint daher als durchaus plausibles Szenario. Der Prozess bis dahin wird sicher nicht mit einem Big Bang kommen, sondern schleichend.

Matt Cutts wies erst kürzlich in einem Video darauf hin, dass Backlinks wohl noch eine ganze Weile ein wichtiger Rankingfaktor bleiben werden. Wenn man die natürliche Sprache besser verstehe, könne man künftig diese Bedeutung immer weiter absenken. Dem Vernehmen nach experimentiert man bei Google immer wieder mit Algorithmen, die Backlinkdaten weitestgehend außen vor lassen – ist aber noch nicht zufrieden mit der Qualität der Ergebnisse. Noch seien die mit gewichteten Linksignalen besser. Noch! Er erwähnte aber auch ganz nebenbei, wenn man besser verstehe, wer etwas geschrieben habe und worum was es genau auf einer Seite gehe, könne man Links durchaus weniger Bedeutung geben. Also doch. ¶

*siehe Online-Glossar unter www.websiteboosting.com