

SEO JANCK hat *PageRank*-Sculpting an drei Nischenprojekten (mit bis zu 100 Seiten), aber wenigen Backlinks, mit „überragendem Erfolg“ eingesetzt (JANCK 2008b). EBACH vermutet eine Ranking-Verbesserung von 0 bis 20% (EBACH 2008a). SEO DANNY SULLIVAN ist allerdings ebenso wie MICHAEL MARTINEZ davon überzeugt, dass *PageRank*-Sculpting keinen sinnvollen Effekt habe: „Leave PageRank Sculpting to idiots who don’t have time to learn about real search engine optimization.“ (MARTINEZ 2008). Laut *seo-theory.com* sei der Einsatz nur bei Seiten sinnvoll, auf denen Besucher Links in eine „Bad Neighbourhood“ setzen könnten, unter denen eventuell die ganze Site leiden könnte – beispielsweise bei Blogs oder Communities (MARTINEZ 2007). Laut *Googles* MATT CUTTS ist *PageRank*-Sculpting ein Effekt zweiter Ordnung (BEUS 2008j). Ursprünglich eingesetzt wurde das Nofollow-Tag, um Links zu Seiten zu kennzeichnen, bei denen man sich nicht sicher war, welche Inhalte sie enthalten. Heute ist es außerdem eine von *Google* propagierte Möglichkeit, Blogs vor Link-Spam schützen.

4.2.3.2 Qualität der verlinkenden Seiten

2003 kaufte *Google* das Patent für den sogenannten Hilltop-Algorithmus und integrierte es vermutlich¹⁰² im Zuge des Florida-Updates im Dezember 2003 in seinen Ranking-Algorithmus. Hintergrund des Hilltop-Algorithmus ist die Annahme, dass der User nach Abbruch einer Suche eine erneute Suche eher auf einer bekannten als auf einer unbekanntem Seite startet. Links von diesen sogenannten „Experten-Seiten“ oder „Authorities“ erhalten deshalb unabhängig vom *PageRank* der verlinkenden Seite eine stärkere Gewichtung. Solche Seiten sind zum Beispiel Webbverzeichnisse wie *Yahoo!* oder das *Open Directory Project*, *Wikipedia* oder andere große Portale (JASTER o.J.). Authorities sind durch folgende Merkmale gekennzeichnet:

- Links zu vielen Seiten zu diesem Thema,
- Inhalte von verschiedenen unabhängigeren Autoren,
- Links von unabhängigen Organisationen.

Durch den *TrustRank*-Algorithmus soll eine automatisierte Unterscheidung zwischen „normalen“ Sites und Spam Sites ermöglicht werden. Der grund-

¹⁰² Mehrheitlich wird davon ausgegangen, dass Hilltop integriert wurde (z. B. FISCHER 2006: 355), letztendlich bleibt die Integration aber Spekulation (vgl. <http://www.suchmaschinen-doktor.de/algorithmen/hilltop.html>).

legende Ansatz des *TrustRank* ist, dass Spam Sites selten oder nie von vertrauenswürdigen Seiten verlinkt werden. Da auch *TrustRank* anhand der Linkstruktur abgeschwächt über mehrere Ebenen vererbt wird, ist es bereits mit einer geringen Auswahl von maximal 200 manuell ausgewählten vertrauenswürdigen Authority Sites möglich sein, Link-Spam wirkungsvoll zu bekämpfen (*Stanford* 2004). CEMPER geht davon aus, dass der „Good Seed“ – also die „gute Saat“ – an vertrauenswürdigen Domains weitaus größer ist; neben einigen 1.000 der über 20.000 .edu-Domains der USA vermutet er auch viele .com- und vielleicht sogar einige .de-Domains (CEMPER 2008b). Das *dmoz*-Verzeichnis beispielsweise gibt zwar laut *Smashing Magazine* keinen extra *PageRank*-Boost weiter, allerdings sehr wohl seinen *TrustRank* (o.V. 2007d). Momentan scheint der *TrustRank* insgesamt wichtiger zum Erreichen einer guten Position zu werden (vgl. *trendmile* 2007b). Andere SEO-Experten gehen sogar davon aus, dass der Trust einer Domain momentan eines der wichtigsten Rankingkriterium ist (GOETZ/TANDLER 2007a). Suchmaschinen gehen zunehmend dazu über, Links von seriösen, als kompetent anerkannten Websites als wichtiges, wenn nicht gar entscheidendes Ranking-Kriterium zu bewerten.

Oft wird davon ausgegangen, dass spezielle Top-Level-Domains (wie .edu oder .gov), welche gesondert beantragt werden müssen, ein höheres Vertrauen von Suchmaschinen entgegengebracht wird als anderen Seiten. Tatsache ist aber, dass nicht primär die Art der Top-Level-Domain für ein besseres Ranking verantwortlich ist, sondern deren Nutzung: So existiert beispielsweise die Domain stanford.edu bereits seit 1985 und besitzt dadurch eine vergleichsweise große und alte Backlink-Struktur. BEUS geht davon aus, dass ein Großteil des Wertes einiger edu-Domains daraus resultieren, dass sie zum „Good Seed“ des *TrustRank*-Algorithmus gehören (BEUS 2008c).

Neben *PageRank* gibt es eine Reihe weiterer, teils bessere Kriterien, um den Wert einer Seite in Bezug auf Backlinks zu beurteilen – Traffic, Anzahl und Qualität der Backlinks und Anzahl der indizierten Seiten sind einige, wie häufig eine Seite gecrawlt und neu indiziert wird, ist ein weiterer.

Auch die Häufigkeit der Aktualisierung einer Webseite kann ein Rankingkriterium sein. Wie häufig sich Inhalte auf einer Webseite ändern, kann beispielsweise mithilfe des Tools *Google Alert* festgestellt werden (vgl. Abschn. 3.3.11). Das Datum des letzten Crawling-Vorganges lässt sich mithilfe des *Google*-Cache ermitteln. Wann die Seite das letzte mal von *Google* indiziert wurde, wird in der erweiterten Suche bei Angabe einer Zeitbegrenzung angezeigt. Da Webseiten direkt nach dem Crawling – dem

eigentlichen Einlesevorgang – indiziert, also ausgewertet werden, kann ein vergleichsweise großer Abstand zwischen beiden Terminen laut HAMLET BATISTA darauf hindeuten, dass *Google* die Seite für wenig wichtig erachtet. Wird eine gecrawlte Seite gar nicht indiziert, kann dies ein Indiz für Duplicate Content sein (*Pole Position Marketing* 2008: 5).

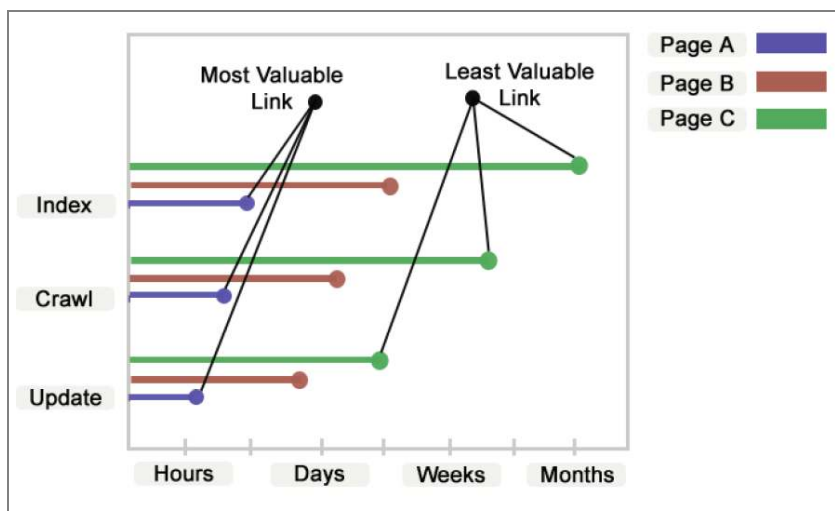


Abb. 60 Crawling-, Indizierungs- und Aktualisierungsfrequenz als Indikator der Qualität (*Pole Position Marketing* 2008: 5)

4.2.3.3 Distanz der verlinkenden Seite

Um bewusste Manipulationen der Linkstruktur – wie zum Beispiel durch „Google Bombings“ – möglichst auszuschließen, werden Quell- und Zielseiten von Links auf ihre Distanz untersucht. Wenn die Distanz zwischen zwei Webseiten zu gering ist, „misstraut“ *Google* unter Umständen der Verlinkung – besonders, wenn davon auszugehen ist, dass beide Webauftritte voneinander abhängig und unter Umständen gar Teil eines Linknetzwerkes sind. Maße für Distanz sind zum Beispiel Domainname, Domaininhaber, Server, IP-Adressen¹⁰³ oder Webautor. BEUS schreibt, dass die aufgrund der IP-Adresse bestimmte Distanz in den allermeisten Fällen kein Rankingfaktor

¹⁰³ Zumindest D-Blöcke der IP-Adresse sollten unterschiedlich sein, damit *Google* davon ausgeht, dass beide Webauftritte unabhängig sind.

ist, da sonst alle, beispielsweise von *Strato* gehosteten Webseiten schlechter ranken würden (BEUS 2008h).

Eine zu große Distanz – beispielsweise durch eine Verlinkung aus unterschiedlichen geographischen Regionen oder von größtenteils anderssprachigen Seiten – kann sich jedoch auch negativ auswirken: Laut Untersuchungen der *c't* war es bereits 2006 nicht mehr möglich, deutsche Seiten durch Links von englischsprachige Seiten mit hohem *PageRank* aufzuwerten – hatte eine Seite zu viele Backlinks von englischen Seiten, wurde sie als englischsprachig eingestuft und in den deutschen SERPs abgewertet (BAGER/BECKER-FOCHLER 2006: 80).

BEUS schreibt 2008, dass *Google* die Zuordnung einer Webseite zum jeweiligen Index eines Landes heute nicht mehr so stark vom Inhalt der Seite als vielmehr von den eingehenden Links abhängig mache. Eine Ausnahme mache *Google* „offenbar“ bei Links von englischen Seiten. Die Auswahl, welcher Sprache eine Seite zugeordnet wird, erfolge dabei nicht grundsätzlich auf Domainebene, sondern vermutlich auf Basis der einzelnen Seiten (BEUS 2008d). Es stellt sich die Frage, ob es beispielsweise möglich ist, eine deutsche Seite mittels vieler Links von chinesischen Seiten aus dem deutschen und in den chinesischen Index zu steuern (vgl. Abschn. 4.4.2).

Auch Verweise, welche aus einem direktem Linktausch resultieren, also von der Zielseite auf die Webseite zurückverlinken, sind ein Indiz für Abhängigkeit und werden vermutlich geringer gewertet.¹⁰⁴

4.2.3.4 Themenbasierte Faktoren der Linkpopularität

Ob die Themenrelevanz der verlinkten Seiten eine Rolle spielt, wird schon lange als mögliches Rankingkriterium diskutiert. Im Jahre 2003 wurde davon ausgegangen, dass es zwar theoretisch möglich, aber technisch noch zu aufwendig wäre, den thematischen Zusammenhang zwischen Seiten valide zu eruieren. Gleichzeitig wurde bezweifelt, ob ein solches Rankingprinzip überhaupt sinnvoll sei, da die Motivation zum Setzen von Links vielfältig ist. Suchmaschinen zum Beispiel würden so nur auf Suchmaschinen verlinken (vgl. *eFactory* 2003: 10).

104 seit dem Bourdon-Update vom Mai 2005

ERLHOFER legt nahe, dass das Ranking seit Januar 2004¹⁰⁵ an eine themenrelevante Verknüpfung gekoppelt ist (ERLHOFER 2006: 356), und laut MEIERT ist die thematische Nähe der verlinkenden Seiten mindestens so wichtig wie die Anzahl der Links (MEIERT 2006). Auch GREIFENEDER sieht den thematischen Zusammenhang zwischen zwei Seiten als Rankingfaktor (GREIFENEDER 2006: 85 ff.).

JANCK geht 2007 davon aus, dass die themenrelevante Verlinkung beachtet wird, bisher aber zumindest abseits der umkämpften Märkte kaum praktische Bedeutung besitze. Ein Nachweis für die Relevanz des thematischen Zusammenhangs kann die Suche nach „Digitalkamera“ versuchen: Die Seite *www.digitalcamera.de* wird bei *Google* trotz wesentlich weniger Incoming Links vor *www.ciao.de* gelistet (JANCK 2007b). BEUS hingegen schreibt, dass die themenrelevante Verlinkung noch keine Rolle spielt, sich dies jedoch in naher Zukunft ändern wird. Als mögliche Gründe werden vermutet, dass der Algorithmus noch nicht so weit entwickelt ist, die Rechenleistung von *Google* zu klein ist oder eine spätere Einführung aus Marketing-sicht besser sein könnte (BEUS 2007k).

Dass der Linktext bereits seit längerem eine wichtige Rolle im Ranking-Algorithmus spielt, zeigt folgendes Beispiel: Wer bei *Google* „hier klicken“ eingibt, bekommt als ersten Treffer die *Adobe*-Seite zum Download des *Acrobat Reader*, auf der dieser Begriff allerdings gar nicht vorkommt. Es gibt jedoch sehr viele Seiten, die mit diesem Link-Text auf die *Adobe*-Seite verlinken. Als dieser Effekt bekannt wurde, entwickelte sich daraus ein „Sport“: Beim sogenannten „Google Bashing“¹⁰⁶ wird versucht, eine Webseite durch Backlinks mit einem Linktext inhaltlich mit diesem zu verbinden. Die Webseite des Weißen Hauses wurde so mittels einer hinreichend hohen Anzahl von Backlinks mit dem Text „miserable failure“ bei *Google* unter diesem Suchbegriff auf Platz eins gelistet.

Im Idealfall sollte der Textinhalt der Backlinks im wörtlichen, zumindest aber in semantischem Zusammenhang mit den relevanten Keywords der Zielseite stehen (vgl. *Bloofusion* 2007: 14). Die Begriffe „hier klicken“ sind demzufolge ein schlechter Linktext. Im Zusammenhang mit dem Linktext werden zudem auch Keywords direkt vor und hinter dem Link beachtet. Laut Black-Hat-SEO RALPH TEGTMEIER ist es sehr sinnvoll, zumindest in

105 *Googles* Austin-Update

106 auch als „*Google Bombing*“ bezeichnet

50 Prozent aller Fälle die Keywords nicht direkt im Linktext, sondern stattdessen in unmittelbarer Nähe zum Link zu positionieren (*LinkJuicy* 2007). BEUS hat im März 2008 bewiesen, dass *Google* nur den jeweils ersten Linktext als Voting für die Zieldomain beachtet, wenn mehrere Links von Site A auf Site B verweisen (BEUS 2008b). Laut EBACH werden bereits seit Februar 2005 Seiten von *Google* entwertet, „deren einkommende Links zum größten Teil aus dem immer gleichen Text bestehen“ (EBACH 2007). Für eine natürlich (aufgebaut) wirkende Backlinklinkstruktur sollte also auch eine gewisse Variabilität der Linktexte vorhanden sein. Dies ist beispielsweise durch unterschiedliche Schreibweisen eines Begriffes oder dessen Nutzung in verschiedenen Kombinationen möglich (vgl. Abschn. 2.7). Tatsache ist auch, dass viele natürlich entstandene Links den Domainnamen anstelle des optimalen Keywords für die Zielseite enthalten. Andererseits zeigt *SEO Scene* an der Seite www.inkandstuff.co.uk (vgl. Abschn. 4.4.1.1 und Abschn. 4.4.1.5), dass Seiten immer noch aufgrund vieler Links von nicht themenrelevanten Seiten mit identischem Linktexten gut ranken können (*nXplorer* 2007).

4.2.3.5 *Format, Position und Aktualität eines Links*

Format und Position eines Links auf einer Webseite sollten nach dem Benutzermodell weitere Faktoren für die Beurteilung von Links sein (vgl. Abschn. 4.2.2). Es ist davon auszugehen, dass ein deutliches Hervorheben eines Links und eine prädestinierte Position innerhalb der einzelnen Seite beziehungsweise der gesamten Webpräsenz dem Ranking der verlinkten Seite eher zugute kommt. Spätestens seit dem *PageRank*-Update vom Oktober 2007 ist zudem erwiesen, dass Links, deren Positionierung und Formatierung auf Linkhandel schließen lässt, zu einer Abwertung der verkaufenden Seite führen können (vgl. Abschn. 4.4 und 4.5) – der Begriff „Content Link“ als Abgrenzung zu Links, welche in weniger wichtigen Bereichen der Seite platziert sind, hat seitdem an Bedeutung gewonnen. Je deutlicher ein Link zu erkennen ist, desto wahrscheinlicher ist es, dass dieser durch einen Besucher verfolgt wird. Links im Footer einer Seite sind tendenziell weniger wichtig als Links im Content. Diesem Gedanken folgend, wäre es möglich, dass Links von einer höheren Navigationsebene – u. U. auch unabhängig vom *PageRank* – bedeutsamer sind als solche von einer tieferen Ebene (JANCK 2007b). Die Textlänge eines Links sollte acht Wörter beziehungsweise 55 Zeichen nicht überschreiten [549]. Nachdem *Google* die ursprüngliche

Link-Popularität durch die Domain-Popularität ersetzt hat, zählt bei mehreren Links von einer Domain „nur noch der Link, der von der Seite mit der höchsten Wertung ausgeht“ (EBACH 2007).

Für eine Suchmaschine ist es von großer Bedeutung, dass die gezeigten Suchergebnisse eine hohe Aktualität besitzen. Demnach muss die Aktualität der Backlinks Einfluss auf das Ranking haben, da bei „kürzlich“ erstellten Backlinks davon ausgegangen werden kann, dass die betreffende Webseite in letzter Zeit von Bedeutung gewesen ist. Andererseits ist es auch wichtig, dass die Linkstruktur über einen längeren Zeitraum natürlich gewachsen ist, also auch ältere Links existieren, da eine unnatürlich rasche Zu- oder Abnahme von Links auf Manipulationen hindeuten könnte. Die Möglichkeiten der Algorithmen, zukünftig sowohl Seitenlayout als auch den Kontext und Nutzung eines Links zu erkennen, werden weiter steigen.

4.2.3.6 Outbound-Links

Outbound-Links verringern entsprechend der *Google*-Formel den Gesamt-*PageRank* einer Webseite, da sie entsprechend dem Zufallssurfer-Modell die Wahrscheinlichkeit erhöhen, dass ein Surfer die Webseite verlässt. Je mehr *PageRank* also nach außen abgegeben wird, desto weniger bleibt innerhalb der Webpräsenz. Da Verlinkung jedoch die Grundlage des Internets bildet, ist davon auszugehen, dass Outbound-Links an anderer Stelle des Algorithmus positiv wirken. Ausgehende Links sollten möglichst auf eine Seite konzentriert werden (*eFactory* 2003: 5). Laut Aussagen von MATT CUTTS verlieren Seiten, die durch Verlinkung „voten“, ihren eigenen *PageRank* nicht, denn: “it’s good for your users, and therefore it’s good for search engines” (SIEMERS 2007f). Im Blog der *Google Webmaster Zentrale* findet sich folgende Aussage: „Sinnvolle ausgehende Links können eure Glaubwürdigkeit erhöhen. Sie zeigen, dass ihr Recherche betrieben habt und Fachwissen auf dem Gebiet besitzt“ [@538]. Am Beispiel von News-Portalen wurde festgestellt, dass Seiten, die weiterführende Links anbieten, zudem stärker frequentiert werden als andere (KARP 2008).

Um den Mehrwert für Besucher auch der Suchmaschine deutlich zu machen, sollte möglichst nur zu qualitativ hochwertigen Seiten verlinkt werden. GREIFENEDER empfiehlt sogar, möglichst nur Links zu Seiten mit einem Mindest-*PageRank* von vier zu setzen (vgl. GREIFENEDER 2006: 81).

Verlinkungen in eine „Bad Neighbourhood“ können hingegen zu einer Abstrafung der eigenen Seite führen.

Andererseits kann auch die Platzierung zu vieler Outgoing-Links auf einer Seite negativ bewertet werden. Laut MATT CUTTS ist es aber mittlerweile unproblematisch, auf einer Seite mit einem hohem *PageRank* „durchaus 200 bis 300 Links“ zu platzieren (*SEOmoz* 2007e). In seinen *Webmaster Guidelines* empfiehlt *Google* maximal 100 Links auf einer Seite zu platzieren [@343]. Einige SEOs vermuten zudem, dass Outgoing Link nur dann positive Effekte haben könnten, solange sie auch Linkjuice weitergeben. No-follow-Links wirkten sich in einem Experiment (von 2005!) vergleichsweise negativ auf das Ranking aus (GÖTTSCHE 2005).

4.2.3.7 Strategien für Verlinkungen

Der ursprüngliche Sinn einer Verlinkung und zugleich die Basis des Internets ist der Verweis zu weiterführenden Informationen. Wenn die Information also bedeutend genug ist, wächst eine entsprechende organische Linkstruktur von selbst. Konsequenterweise bewertet *Google* letztendlich übertriebene Versuche, sich proaktiv um Verlinkung zu bemühen, als Manipulation. Anfang August 2007 hat *Google* in seinen *Richtlinien für Webmaster* entsprechende Passagen eingefügt und setzt diese seit Ende August 2007 auch durch (vgl. Abschn. 4.2.5.3). Möglichkeiten zur Manipulation, von denen abgeraten wird, sind:

- Links mit der Absicht, den *PageRank* zu manipulieren,
- Links in eine „Bad Neighborhood“,
- Ankauf und Verkauf von Links, die *PageRank* vererben,
- exzessive reziproke Verlinkung.

Die Grenze zwischen Optimierung und Manipulation ist dabei fließend. Egal, durch welche Methoden der Einfluss des Link-Wachstums forciert wird, entscheidend ist, dass es natürlich wirkt. Dies gilt sowohl für die Geschwindigkeit des Linkwachstums, als auch für die Menge und Position der Links (WARD 2007a) (vgl. Abschn. 4.2.5.4). Die PR-Struktur aller Backlinks sollte weitestgehend eine Pyramidenform aufweisen.

TEGTMEIER empfiehlt in diesem Zusammenhang, die Mehrzahl aller Links auf die Home-Seite verweisen zu lassen, jedoch auch eine kleinere Menge Deep Links auf Unterseiten zu setzen (*LinkJuicy* 2007).

Finden und Beurteilen von Linkquellen

Besonders im Zusammenhang mit Linkkauf und ist es wichtig, die Wertigkeit eines Links beurteilen zu können. Links sollten immer als normale HTML-Links genutzt werden. Weiterleitungen mittels verschiedener Scripting-Techniken können eine Weitergabe des *PageRank* genauso verhindern wie der Einsatz des *Nofollow*-Tags, welcher direkt im Link oder global für die Webseite definiert werden kann. Außerdem muss die verlinkende Seite im Index der Webseite vorhanden sein. Besonders geeignet für eine Verlinkung sind Webseiten, die:

- für das gewünschte Keyword sehr weit vorn ranken,
- Authorities sind, also neben diesem auch für andere Keywords weit vorn ranken.

Für wie wichtig *Google* eine Seite hält, kann ungefähr am dargestellten *PageRank* (TBPR) abgelesen werden – *Google* nutzt diesen Wert nach eigenen Aussagen mittlerweile als eine Art Kommunikationsinstrument (vgl. Abschn. 3.1.3.3). Bis April 2008 wurde der – im *Google Directory* – dargestellte *PageRank*-Wert auf einer Siebenerskala dargestellt. Mittlerweile wurde aber auch diese Skala auf Zehnerschritte umgestellt. Wird der PR-Wert einer Seite nicht angezeigt, kann dies durchaus auch daran liegen, dass sie so neu ist, dass *Google* ihr (extern) noch keinen *PageRank* sichtbar zugewiesen hat.

Zur Bewertung der Wichtigkeit von Webseiten kann auch der *Alexa*-Rank genutzt werden. Das *Alexa*-Ranking basiert auf der Anzahl Besucher, welche die *Alexa*-Toolbar¹⁰⁷ installiert haben. Auf der *Alexa*-Webseite lassen sich Reichweite, Ranking und Page Views mehrerer Seiten anhand der von *Alexa* ermittelten Daten über Zeiträume vergleichen. Im April 2008 wurde der zugrunde liegende Ranking-Algorithmus angepasst, um die Vergleichbarkeit der Daten zu verbessern: Seiten mit vielen *Alexa*-Zugriffen wurden im Vergleich deutlich abgewertet. So hat *Yahoo!* beispielsweise ein Ranking von 1, *Google* von 17 (vorher: 2) und die Seite der HTWM¹⁰⁸ von 99.230 (vorher: 164.542). Trotzdem sollten die *Alexa*-Rankings kritisch betrachtet werden, da die Toolbar bis August 2007 nur Nutzern des *Internet Explorers* zur Verfügung stand, eher in den USA verbreitet ist und die Nutzer meist nur die eher ambitionierten Surfer sind. Für einen brancheninternen Vergleich von

107 Download unter: <http://www.alexa.com/site/download/>

108 Hochschule für Technik und Wirtschaft Mittweida (FH)