

Masterseminar MI-Sem-M — Wintersemester 2010/11

zum Thema

Enterprise Search

Otto-Friedrich Universität Bamberg
Lehrstuhl für Medieninformatik
www.uni-bamberg.de/minf

Seminararbeit zum Thema

Worst Practices in SEO

Michael Fischer

mail@michaelfischer.name

Zusammenfassung Diese Arbeit befasst sich mit Blackhat- und Spamming Methoden sowie der inkorrekten Suchmaschinenoptimierung. Anhand einer Internetrecherche wurden verschiedene Worst Practises ausfindig gemacht. Kategorisierungs- und Rankingvorschläge führen zu einer strukturierten Übersicht der Praktiken. Weiterhin werden die Praktiken mit den negativsten Folgen für Suchmaschinen detailliert untersucht. Hierbei lassen sich die Einflüsse auf Suchmaschinen erklären und Lösungsmöglichkeiten aufzeigen. Zum Abschluss behandelt die Ausarbeitung mögliche Bestrafungen und zeigt weitere Vertiefungsmöglichkeiten auf.

Inhaltsverzeichnis

1	Einleitung	3
1.1	Motivation	3
1.2	Definitionen	4
2	Vorgehensweise	6
3	Worst Practices im Überblick	7
3.1	Möglichkeiten einer Kategorisierung	7
3.2	Möglichkeiten einer Rangliste	8
4	Worst Practices im Detail	8
4.1	Improper Keyword Management	8
4.2	Cloaking	10
4.3	Content Duplication	12
4.4	Massively Linking from Outside	13
4.5	Poor Coding and Design	14
5	Schlusswort	16
5.1	Mangelnde Suchmaschinenoptimierung	16
5.2	Blackhat Methoden und Suchmaschinen-Spamming	17
5.3	Angewandte Filter der Suchmaschinen	18
5.4	Weitere Themengebiete	18
6	Anhang	21

1 Einleitung

Diese Arbeit beginnt mit einer Hinführung zu dem Thema SEO Worst Practices, gefolgt von einer Abgrenzung anhand von Definitionen. Dadurch wird verdeutlicht, welche Aspekte innerhalb dieser Ausarbeitung betrachtet werden. Am Ende der Einleitung wird der Bezug zu Enterprise Search erläutert, da diese Arbeit im Rahmen eines Seminars zum Thema Enterprise Search entstanden ist.

1.1 Motivation

Das Interesse an Suchmaschinenoptimierung ist in den letzten Jahren stetig gestiegen ¹. Hierbei entstand eine Fülle an Literatur für positive Suchmaschinenoptimierung und jede Menge Best Practices. Zu Worst Practices innerhalb der Suchmaschinenoptimierung gibt es jedoch sehr wenig Literatur, besonders auf wissenschaftlicher Ebene. Diese Lücke soll mit dieser Ausarbeitung geschlossen werden. In dieser Arbeit werden die Techniken, welche hinter Worst Practices stehen, aufgezeigt und einige von Ihnen näher erläutert.

Worst Practices können aus verschiedenen Absichten heraus entstehen. Es ergeben sich hierbei unterschiedliche Ansätze: Unter einer Absicht versteht man das bewusste Anwenden von Blackhat Methoden oder Suchmaschinen-Spamming. Der Versuch, das eigene Ranking durch unzulässige oder grenzwertige Techniken zu verbessern. Diese können bei Erfolg zu positiven Ergebnissen führen. Werden solche Methoden jedoch erkannt oder zu massiv verwendet, führen sie zu negativen Ergebnissen. Zum Beispiel hatten vor einigen Jahren namhafte Unternehmen wie BMW sowie die Gebrauchtwagenplattform automobile.de unerlaubte Methoden angewandt und wurden deshalb für einige Zeit aus dem Index entfernt [1]. Eine andere Absicht legt besonderen Wert auf die Webseiten der Konkurrenten. Diese werden durch den bewusst übertriebenen Einsatz von Blackhat und Spamming Methoden angegriffen mit dem Ziel, die Webseite des Mitstreiters aus den Suchergebnissen verschwinden zu lassen. Häufig wird der Begriff *Google Bowling* ² oder *Reverse SEO* ³ dafür verwendet. Es gibt sogar Anbieter, welche Blackhat Methoden bei Webseiten gegen eine entsprechende Bezahlung durchführen [2]. Als dritter Ansatz gilt das unbewusste Anwenden von Worst Practices, welches zu einem schlechten Ranking der eigenen Seite führt. Hierbei wird keine besondere Absicht verfolgt, sondern Hauptgründe sind fehlendes Wissen für die Gestaltung einer suchmaschinenoptimierten Seite oder schlampige Durchführung der Seitenerstellung. Meistens können Suchmaschinen in diesem Fall die Seiten nur schlecht durchsuchen oder die relevanten Schlüsselbegriffe nicht finden.

Im Folgenden wird ein Beispiel [3] etwas detaillierter beschrieben, welches am ehesten dem Blackhat Bereich zugeordnet werden kann. In diesem Fall versucht

¹ vgl. <http://www.google.com/insights/search/#content=1&cat=13&q=seo&cmpt=q> (Abruf: 26.3.11)

² vgl. <http://seotermglossary.com/google-bowling/> (Abruf: 26.3.11)

³ vgl. <http://www.searchenginepartner.com/First-page-placement/reverse-seo.html> (Abruf: 26.3.11)

ein Onlineshopbetreiber besonders schlecht mit seinen Kunden umzugehen. Mit Absicht liefert er Mangelware und droht seinen Kunden gleichzeitig an, er solle keinesfalls die Ware zurückschicken oder das bereits bezahlte Geld zurückfordern. Die Aktionen von Vitaly Borker, dem Webseitenbetreiber, sind aus folgendem Grund bewusst inszeniert: Suchmaschinen können nur sehr schwer zwischen positiv, neutral oder negativ behafteten Links unterscheiden. Aufgrund der enormen Beschwerden von Kunden wurden durch diese automatisch Links zu seiner Seite generiert. Dadurch stieg der Wert seiner Seite und diese wurde bei Suchergebnissen höher bewertet. Folgerichtig lässt sich feststellen, wer sich im E-Commerce übermäßig schlecht verhält, kann durch diese Methode erhöhte Gewinne realisieren. Die Reaktion der Suchmaschinen folgte prompt. Folgende Ansätze wurden von der Suchmaschine Google als Lösungsmöglichkeit genannt [4]:

Das individuelle Blockieren des speziellen Täters handelt von einer sofortigen Lösung des speziellen Falles. Suchmaschinen legen jedoch einen hohen Wert auf eine automatisierte Lösung des Problems, bei der nicht in jedem Einzelfall neu gehandelt werden muss.

Als zweite Möglichkeit gilt die Benutzung von Sentiment Analyse. Ein guter Ansatz, welcher mit zwei Problemen versehen ist. Neben den generierten Links mit negativer Resonanz, gibt es ebenso viele Berichte von reputablen und populären Webseiten, besonders Nachrichtenportalen, bei denen eine Sentimentanalyse nicht anwendbar ist. Ferner könnten kontroverse Diskussionen oder Informationen zu bestimmten Personenkreisen (zum Beispiel Politiker) in den Suchergebnissen verschwinden.

Bei der Anzeige von Benutzerkritiken neben den eigentlichen Ergebnissen werden die Webseiten trotzdem gut bewertet. Zusätzlich ist es für Suchmaschinen schwer zu entscheiden, bei welchen Seiten die Benutzerkritiken erscheinen sollen.

Als Lösung wird wahrscheinlich ein Algorithmus verwendet, welcher die oben genannten Lösungsmöglichkeiten mit einbezieht. Über die verwendeten Algorithmen geben die Suchmaschinenbetreiber selbstverständlich keine Auskunft. Ziel ist es Betreiber mit schlechten Nutzererfahrungswerten zu entdecken. Diese Methode, bei der ein Schlupfloch im Ranking auf sehr radikale und teilweise illegale Weise ausgenutzt wird, stellt eher eine Ausnahme dar. In dieser Ausarbeitung werden besonders die technischen Aspekte der zuvor erwähnten Ansätze von Worst Practices in SEO beleuchtet.

1.2 Definitionen

SEO:

Im Deutschen steht SEO für Suchmaschinenoptimierung. Jaimie Sirovich und Cristian Darie, zwei anerkannte Programmierer mit besonderem Bezug zu Suchmaschinenmarketing, verstehen darunter das Ziel zu erlangen, die Anzahl der Besucher einer Internetseite durch das Verbessern von Positionen von unbezahlten und organischen Suchmaschinenergebnissen zu erhöhen [5]. Hierbei handelt es sich besonders um eine technische und inhaltliche Anpassung der eigenen Seite, sowie das richtige Verknüpfen mit anderen Seiten. Als Gegenpart hierzu dient das Suchmaschinenmarketing, welches sich mit dem gezielten Einsatz von

bezahlten Einträgen beschäftigt [6]. Dieser Part wird innerhalb dieser Ausarbeitung nicht betrachtet. Der Begriff Suchmaschinenmarketing wird teilweise auch als Oberbegriff für Suchmaschinenoptimierung und bezahlte Einträge verwendet.

Zukünftig wird mobile SEO oder mobile SEM einen höheren Einfluss haben. Die steigende Anzahl an mobilen Internetnutzern ist ein wichtiges Indiz hierfür: Die Zahl der Internetnutzer, welche sich über das Handy einwählten, stieg nach Angaben des statistischen Bundesamtes im Jahre 2010 um 78% im Vergleich zum Vorjahr auf 17%. Deutschland steht statistisch hinter den Nachbarländern Österreich und Schweiz, die ungefähr 10% höher liegen [7]. Diese Ausarbeitung bezieht sich ausnahmslos auf SEO und Worst Practices und nimmt keinen Bezug zu mobile SEO oder SEM.

Worst Practices:

Der Begriff wird häufig mithilfe des Begriffs Best Practices hergeleitet. Diese handeln von einem unternehmensübergreifenden Einsatz von herausragenden, exzellenten Praktiken, um eine nachhaltige Verbesserung oder einen Wettbewerbsvorteil zu erlangen [8]. Unter Worst Practices wird das Gegenteil verstanden: Sie gelten als ineffizient, unproduktiv und führen häufig nicht zu dem Erfolg eines speziellen Zieles. Worst Practices werden häufig nicht kommuniziert oder veröffentlicht, um ein negatives Bild vom Scheitern seines eigenen Unternehmens zu vermeiden. Aus diesem Grund gibt es sehr wenig Literatur zu Worst Practices im Vergleich zu Best Practices. Das Lernen aus Worst Practices und die Vermeidung dieser kann wiederum zu Best Practices führen ⁴.

SEO Worst Practices:

Bei der Kombination beider Begriffe ergeben sich drei verschiedene Methoden, welche am häufigsten unter der Suche nach den Begriffen von SEO und Worst Practices aufgetreten sind: Diese Methoden sind vergleichbar mit den Absichten, welche bereits in der Motivation erläutert wurden.

Unter Blackhat SEO werden Verletzungen der Richtlinien von Suchmaschinen bewusst in Kauf genommen. Massive Spamming-Praktiken, wie zum Beispiel *Cloaking* oder *Doorway Pages*, werden eingesetzt um das Ranking zu verbessern. Diese Techniken können als Worst Practices angesehen werden, da davon ausgegangen wird, dass Suchmaschinen Blackhat Techniken innerhalb kurzer Zeit entdecken. Demnach ist eine nachhaltige Verbesserung nicht möglich. Im Gegenteil, diese Anwendung führt sehr häufig zu negativen Einflüssen [9].

Suchmaschinen-Spamming, als zweite Möglichkeit, tritt meistens ebenso bewusst auf. Es ist die Rede von einer Überoptimierung der eigenen Seite oder einem exzessiven Gebrauch von Suchmaschinenoptimierungstechniken wie zum Beispiel *Link Spamming*. Hierbei wird versucht die Grenze des erlaubten Handelns innerhalb Suchmaschinenoptimierung so weit wie möglich auszuschöpfen. Häufig wird die Grenze überschritten und als Folge strafen Suchmaschinen diese Seiten ab.

⁴ vgl. http://de.wikipedia.org/wiki/Best_Practice (Abruf: 26.3.11)

Ungenügendes Wissen über Suchmaschinenoptimierung und deren Best Practices führt zu schlechten Positionierungen innerhalb von Suchmaschinenergebnissen. Die Mitbewerber können sich einen Wettbewerbsvorteil verschaffen durch die Anwendung von Suchmaschinenoptimierung. Hierzu zählen Techniken, die dem Begriff *Poor Coding and Design* zuzuordnen sind. Beispiele sind Seiten mit vielen Popups, zu vielen Bildern, zu wenig Textinhalt, etc...

Bezug zu Enterprise Search:

Diese Ausarbeitung entstand innerhalb des Seminars *Enterprise Search* am Lehrstuhl für Medieninformatik an der Otto-Friedrich Universität Bamberg. Das Thema *Worst Practices in SEO* wurde als Sonderthema behandelt. Folgende Aussage nimmt eine Abgrenzung zwischen Enterprise Search und SEO vor.

Während sich Enterprise Search mit Information Retrieval innerhalb Unternehmen befasst [10, S. 333], bezieht sich SEO in der allgemein verwendeten Literatur auf das Internet (siehe SEO Definition). SEO kann als Teil von Information Retrieval verstanden werden. Im Information Retrieval treten Unterschiede zwischen einer unternehmensweiten und einer globalen Suche auf. Innerhalb Unternehmen kommt es darauf an, die relevanten Dokumente zu finden. Im Internet dagegen liegt ein erhöhter Fokus auf einer Rangliste von vielen relevanten Dokumenten. Während im Internet die Linkstruktur für das Auffinden von Dokumenten sehr wichtig ist, wird im Enterprise Search eine Navigation über die Verzeichnisstruktur bevorzugt. Zusätzlich ist die Dokumentvielfalt unterschiedlich [11, S. 2]. Aus diesen Unterschieden lässt sich folgern, dass eine Übertragung von SEO innerhalb des Internets auf SEO innerhalb Unternehmen genau untersucht werden muss. Diese Ausarbeitung behandelt ausschließlich SEO im Internet. Im Unterkapitel 5.4 werden zu untersuchende Fragestellungen bezüglich SEO und Enterprise Search aufgegriffen.

2 Vorgehensweise

Um eine aussagekräftige Ausarbeitung zu schreiben, wurde zuerst nach Worst Practices im Bereich Suchmaschinenoptimierung gesucht (siehe Tabelle 4). Als Suchbegriffe dienen die vorher definierten Begriffe. Da es sehr wenige relevante Artikel und Bücher zu Worst Practices gibt, sind sehr viele Internetseiten mit herangezogen worden. Dabei sind folgende Probleme aufgetreten: Die Begriffe waren teilweise Synonyme, mehrdeutig und in verschiedenen Abstraktionsebenen vorhanden. Weiterhin gab es keine durchgängige Kategorisierung der Begriffe. Zusätzlich wiesen die Artikel eine hohe Korrelation untereinander auf. Sofern zwei bis drei gleiche Begriffe auftraten, waren die gelesenen, untersuchten Artikel meist sehr ähnlich. Folglich ist davon auszugehen, dass einige Artikel voneinander abgeschrieben wurden. Für eine Strukturierung wurden zunächst Kategorisierungsmöglichkeiten (siehe Tabelle 6) untersucht sowie Begriffe zusammengefasst. Beispielhaft wurde eine mögliche Rangordnung der Worst Practices erzeugt (siehe Tabelle 4). Anschließend wurden ähnliche Methoden sowie Methoden mit ähnlichen Auswirkungen zusammengefasst. Die wichtigsten werden im Kapitel 4

näher beschrieben. Hierbei wird auf eine kurze Beschreibung, Definition, Anwendungsbeispiele, die folgenden Konsequenzen und Lösungsmöglichkeiten oder Vermeidungsaspekte eingegangen.

3 Worst Practices im Überblick



Abbildung 1. Auswahl an Worst-Practices dargestellt in einer Tag-Cloud

Bei meiner Suche nach Worst Practices im Bereich SEO ergaben sich 68 unterschiedliche Begriffe (siehe Tabelle 4). Diese entstanden aus 14 verschiedenen Quellen (siehe Tabelle 5). Im Folgenden gehe ich auf Möglichkeiten der Kategorisierung dieser Begriffe ein und zeige eine initiale Kategorisierung sowie eine Platzierung auf.

3.1 Möglichkeiten einer Kategorisierung

Die Begriffe können gruppiert werden in *Onpage* und *Offpage* sowie beabsichtigte und unbeabsichtigte Methoden. Die Gruppierung in *Onpage* und *Offpage* Methoden innerhalb SEO lässt sich auf Worst Practices übertragen. *Onpage* Techniken handeln von Worst Practices bezogen auf die eigene Seite, während *Offpage* Techniken maßgeblich Einfluss auf die Bewertung der eigenen Seite durch andere Webseiten nehmen. Die zweite Gruppierungsmöglichkeit wurde bereits in der Einleitung beschrieben. Zusätzlich können die Begriffe zusammengefasst und kategorisiert werden, indem die Vorgehensweisen oder der gemeinschaftliche Bezug

zu Suchmaschinenoptimierung verglichen werden. Als Oberbegriffe für Kategorien eignen sich zum Beispiel: *Cloaking*, *Keyword Related*, *Link Related*, *Content Related*, *Index Related*, *Poor Coding & Design*. Im Anhang befindet sich eine mögliche Aufteilung der gefundenen Begriffe (siehe Tabelle 6). Es gibt weitere Begriffe, welche diesen Gruppen nicht zugeteilt werden konnten. Diese wurden unter der Kategorie *Others* geführt.

3.2 Möglichkeiten einer Rangliste

Anhand der gefundenen Worst Practices wurde ein initiales Ranking erstellt. Dabei wurden qualitativ höher empfundene Quellen zweifach gewertet, die restlichen einfach (siehe Tabelle 5). Das Ergebnis ist unter der Tabelle 4 zu sehen. Danach wurden ähnliche Methoden - im Bezug auf ihre Vorgehensweise oder Auswirkungen - zusammengefasst, damit mehrere Aspekte innerhalb dieser Ausarbeitung betrachtet werden können. Die am negativ belasteten Worst Practices werden im Nachhinein vorgestellt und erläutert. Dabei wird zuerst eine Definition angegeben, danach auf den Einfluss auf Suchmaschinen eingegangen und zuletzt werden Lösungs- oder Vermeidungsmöglichkeiten vorgeschlagen. Für die Gewichtung der Worst Practices wurden folgende Aspekte mit einbezogen: Die Anzahl der Vorkommen in der Artikelauswahl, die Gewichtung der Methoden innerhalb der Artikel, das Ausmaß des negativen Einflusses auf Suchmaschinen und eine subjektive Einschätzung der Methoden. Aufgrund der Themenstellung innerhalb einer Seminararbeit, wird die Methodik nicht detaillierter beschrieben. Nach der Zusammenfassung der gerankten Worst Practices ergaben sich folgende Methoden:

4 Worst Practices im Detail

Im Folgenden werden die fünf negativsten Methoden der Reihenfolge nach, beginnend mit den niedrigsten Positionen, erläutert.

4.1 Improper Keyword Management

Definition:

In diesem Fall treten viele ähnliche Verfahren auf, bei denen Schlüsselbegriffe suboptimal verwaltet werden. Schlüsselbegriffe sollten ein Vorkommen zwischen 3 und 7% im Verhältnis zu dem kompletten Text aufweisen. Bei einem Vorkommen unter 3% werden Sie nicht als Schlüsselbegriff angesehen. Bei Werten über 7% spricht man von *Keyword Stuffing*, die Schlüsselbegriffe treten zu häufig auf. Solche Sätze sind von Menschen kaum lesbar und ergeben wenig Sinn. Im Folgenden werden einige Methoden kurz beschrieben, deren Einflüsse auf Suchmaschinen und Lösungsmöglichkeiten sehr ähnlich sind:

Mit *Keyword Stuffing* versteht man das übermäßige Beladen von Inhalt oder Meta Tags einer Seite mit Schlüsselbegriffen ⁵. Bei *Keyword Dilution* werden zu

⁵ vgl. <http://www.webconfs.com/over-optimization-article-8.php> (Abruf: 26.3.11)

Tabelle 1. Mögliche Rangliste von gruppierten Methoden

P.	A.	Hyperonym	Methoden
1	31	Poor Coding & Design	Wrong Redirects, Use of Frames/Popups, Dynamic/Unreadable URLs, ...
2	27	Massively Linking from Outside	Link farms, Paid Links, Cross Linking, Link Spamming (Guestbook, Wikis, Blogs, Referrer Log), Free For All Linking Pages, Reciprocal Link Trading, Google Bombing, HTML Injections, Scross Site Scripting
3	20	Content Duplication	Content Generators, Illegal Content, Pagejacking, Article Spinning, Multiple Domains, Scraper Sites
4	22	Cloaking	Doorway Pages, Content Separation, Dorrway Domains, Meta Refresh, Invisible Elements
5	15	Improper Keyword Management	Keyword Stuffing, Keyword Dilution, Splash Page, Superfluous Text, Irrelevant Keywords, Spamglish, Page Dedication to Unusable Keywords

P. = Position im Ranking

A. = Anzahl der Vorkommen der zusammengehörigen Methoden

viele Schlüsselbegriffe gewählt und versucht zu optimieren. Als Ergebnis wird meistens kein einziger Schlüsselbegriff höher bewertet. Als Faustregel gilt: Die Anzahl der Seiten pro Homepage entspricht der Anzahl der maximalen verschiedenen Schlüsselbegriffe, die verwendet werden sollen. Als *Splash Page* wird eine Seite bezeichnet, die wenig textlichen Inhalt bietet, stattdessen ausschließlich Grafiken, animierte Effekte oder Videos. Als Beispiel dienen Introseiten, bei denen zusätzlich die Links nicht in reiner HTML Form, sondern in Javascript angegeben sind. Suchmaschinen erfassen diese Seiten nur sehr schwer. Unter *Superfluous Text* versteht man Texte wie "Herzlich Willkommen auf unserer Homepage". Diese sind nicht themenrelevant und überflüssig. Sie verfälschen bzw. schwächen die eigentlich relevanten Begriffe. *Irrelevant Keywords* sind falsche oder irrelevant verwendete Schlüsselbegriffe. Irrelevante Schlüsselbegriffe erhalten meist sehr gute Positionierungen, sie werden aber kaum gesucht. Bei falschen Begriffen passen die Begriffe nicht zu der eigentlichen Thematik der Homepage. Dies führt häufig zu einem schnellen Verlassen der Internetseite des Besuchers.

Einfluss auf Suchmaschinen:

Der Einfluss wird am Beispiel von *Keyword Stuffing* verdeutlicht. Geht eine Suchmaschine von *Keyword Stuffing* aus, wird im besten Fall nur das Ergebnis des betroffenen Schlüsselbegriffs in der Rangliste heruntergesetzt. Im schlimmsten Fall führt *Keyword Stuffing* zu einer temporären oder permanenten Sperre der kompletten Seite. Dies kann passieren, wenn Unternehmen oder Webseitenbetreiber bewusst auf jeder einzelnen Seite *Keyword Stuffing* betreiben und der Grenzwert von 7% häufig überschritten wird.

Lösungsmöglichkeiten:

Es gibt sogenannte *Keyword Density Tools* ⁶⁷. Diese können genutzt werden um die Verteilung der Begriffe einer Internetseite ausfindig zu machen. Häufig entfernen diese Tools *Noise* (Tag-Wörter, Werbeblöcke, Navigation, ...) automatisch, so wie es auch Suchmaschinen machen. Begriffe mit zu hoher Verteilung müssen demnach sparsamer verwendet werden und vice versa. Sofern eine Seite bereits als *Keyword Stuffing* gekennzeichnet ist, sollte man versuchen diese Seite anschließend bei den Suchmaschinen erneut vorzuschlagen [12].

4.2 Cloaking**Definition:**

Cloaking bedeutet der Suchmaschine eine optimierte Seite für die Indexierung auszuliefern. Nur die personellen Besucher der Internetseite erhalten die eigentliche Seite ⁸. Bei einer Anfrage der Webseite wird bestimmt ob es sich um einen menschlichen Besucher oder um einen *Spider* oder *Crawler* handelt. Für letztere wird eine optimierte Seite bezüglich der Begriffe vorgehalten. Bei menschlichen Besuchern leitet man direkt zu der eigentlichen Webseite weiter. Die Weiterleitung geschieht in so kurzer Zeit, dass sie von den Besuchern nicht wahrgenommen wird [13, S. 2].

Ähnliche Methoden:

Als ähnliche Methode gilt das Verstecken von Elementen. Zum Beispiel kann Text vor dem Besucher versteckt werden, indem die Textfarbe der Hintergrundfarbe entspricht. Diese Methode dient nur als einfaches Exemplar, sie wird von den gängigen Suchmaschinen erkannt.

Ebenso ähnlich wie *Cloaking* sind *Doorway Pages* (siehe Abbild 2). Bei *Doorway Pages* wird eine optimierte Seite für Suchmaschinen vorgeschaltet, welche durch eine zügige Weiterleitung auf eine andere Seite für den Besucher nicht sichtbar ist. Es gibt drei Möglichkeiten eine versteckte Weiterleitung einzurichten: Die Eintragung eines *Meta Refresh* in den *Meta Tags*, Weiterleitungen durch JavaScript oder das Nutzen von serverseitigen Programmiersprachen (zum Beispiel eine Weiterleitung mit PHP definieren) [13, S. 3].

Doorway Domains übertragen das Konzept der *Doorway Pages* auf die Ebene von Domains. Diese Domains gelten als Traffic-Fänger, Zugang und Popularitätssteigerer für die eigentliche Domain ⁹.

Einfluss auf Suchmaschinen:

Sofern Suchmaschinen *Cloaking* auffinden und dies als Regelverstoß bewerten, werden die betreffenden Seiten meistens komplett aus dem Index entfernt. Dies

⁶ vgl. <http://webconfs.com/keyword-density-checker.php> (Abruf: 26.3.11)

⁷ vgl. <http://www.linkvendor.com/seo-tools/keyword-density.html> (Abruf: 26.3.11)

⁸ vgl. <http://www.webopedia.com/TERM/C/cloaking.html> (Abruf: 26.3.11)

⁹ vgl. <http://www.sumo.de/definition-doorway-domain.html> (Abruf: 26.3.11)

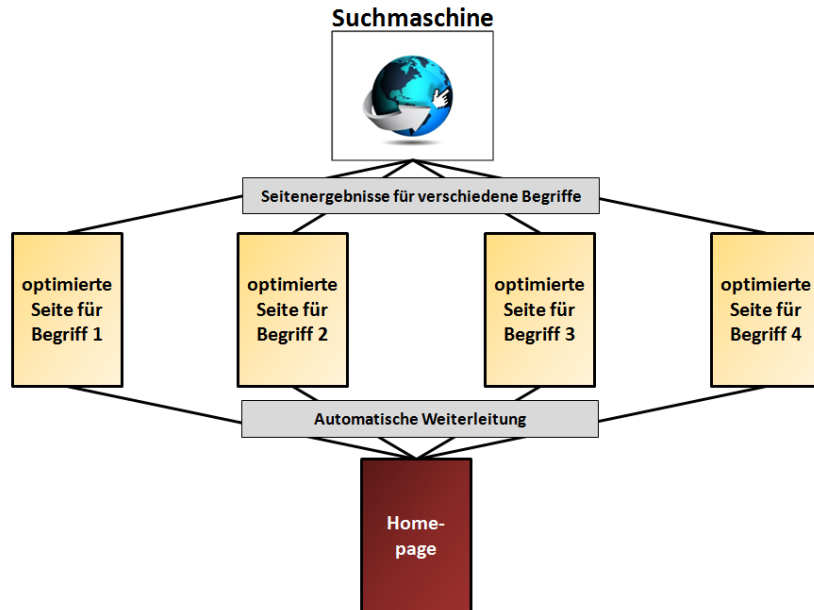


Abbildung 2. Die Vorgehensweise bei Doorway Pages

kann letztendlich zu einer permanenten Sperre führen. Suchmaschinen haben ebenso generelle Lösungen gefunden, um *Cloaking* abzuschwächen. Bei Suchmaschinenenergebnissen zu einer bestimmten Anfrage werden von einer Seite nicht mehr als zwei oder drei Ergebnisse angezeigt. Dadurch macht es keinen Sinn, verschiedene optimierte Begriffseiten zu entwickeln, welche alle auf eine Inhaltsseite zurückzuführen sind. Zusätzlich wurde der Fokus auf eingehende Links verstärkt. Die optimierten Seiten für Suchmaschinen weisen in der Regel kaum eingehende Links auf, da Besucher diese nie zu Gesicht bekommen.

Im Jahre 2006 wurde die Webseite des Automobilherstellers BMW aus den Suchergebnissen von einer der bekanntesten Suchmaschinen genommen. BMW wurde vorgeworfen, *Doorway Pages* zu verwenden. Aufgrund der Größe des Unternehmens waren die Ergebnisse nach ein paar Tagen wieder sichtbar. Viele Nachrichtenportale veröffentlichten diesen Vorfall, sodass BMW dies zu einer Art Marketing Gag transformierte [14].

Lösungsmöglichkeiten:

Grundsätzlich wird von der Verwendung von *Cloaking* abgeraten [13, S. 4]. Es gibt jedoch einige wenige Ausnahmen, bei denen diese Technik angewendet werden darf. Es ist dann insbesondere darauf zu achten, *Cloaking* verantwortungsvoll einzusetzen. Die Schlüsselbegriffe und der Text sollten genau zu der weitergeleiteten Seite passen. Als Beispiel dient eine Flash-Seite, bei der Suchmaschinen

keinen textlichen Inhalt extrahieren können. Dort wird *Cloaking* als Hilfsmaßnahme angesehen und akzeptiert.

Webseitenbetreiber können ebenso passiv zum Opfer von *Cloaking* werden, indem andere Personen versuchen mithilfe von *Cloaking* die Seite des Webseitenbetreibers aus den Suchergebnissen zu entfernen. In diesem Fall sollte der Webseitenbetreiber diesen Vorgang schnellstmöglich den Suchmaschinen melden, um eine Bestrafung vorzubeugen.

4.3 Content Duplication

Definition:

Unter Inhaltsduplikaten versteht man Inhaltsbereiche, welche anderem Inhalt gleichen oder in hohem Maße ähneln. Inhaltsduplikate treten innerhalb Domains oder domainübergreifend auf, unbeabsichtigt oder beabsichtigt: Zu den unbeabsichtigten Methoden gehören gewöhnliche Artikelverteilungen, E-Commerce Produktbeschreibungen sowie Seiten, die für das Web und Mobilgeräte optimiert sind. Zu den beabsichtigten Methoden zählen *Pagejacking*, *Article Spinning*, *Scraper Sites* und Inhaltsgeneratoren. Übersetzungen sowie *Snippets* werden für gewöhnlich nicht als Duplikate angesehen ¹⁰.

Einfluss auf Suchmaschinen:

Suchmaschinen gehen grundsätzlich von unbeabsichtigten Inhaltsduplikaten aus, weil es unabwendbar ist, dass jeder Inhalt nur einmalig vorhanden ist. Zum Beispiel werden gleiche Produkte auf mehreren Internetseiten angeboten. Die Produktdaten können hierbei aus einer gemeinsamen XML Datei automatisch generiert werden. Dies führt zu gleichen Produktbeschreibungen auf mehreren Seiten. Bei Suchmaschinen werden die Webseiten mit diesen Duplikaten nicht abgestraft. Domainübergreifend kann es vorkommen, dass bei Produktbeschreibungen alle Seiten angezeigt werden. Bei Artikelverteilungen wird ein Filter angewendet, welcher wahrscheinlich nur die populärste Seite mit dem Artikel anzeigt. Im Allgemeinen lässt sich schließen, dass es bei unabsichtlichen Inhaltsduplikaten zu einer Filterung kommt, die Duplikate entfernt und nur eine Seite anzeigt [15]. Sofern Suchmaschinen beabsichtigte Methoden erkennen, kann es zu einer Abstufung der Ergebnisse oder einem Entfernen einer kompletten Webseite bzw. Domain für eine längere Zeit führen.

Lösungsmöglichkeiten:

Es gibt Richtlinien, an welche sich Nutzer halten sollten um Inhaltsduplikate zu vermeiden [15]. Werden diese Richtlinien eingehalten, besteht keine Gefahr aus den Suchergebnissen gesperrt zu werden. Gelegentlich kann es vorkommen, dass Seiten, welche domainübergreifend oder innerhalb der Domain nicht einzigartig sind, herausgefiltert werden. Wer auf Nummer sicher gehen möchte, hat die

¹⁰ vgl. <http://www.webconfs.com/duplicate-content-filter-article-1.php> (Abruf: 26.3.11)

Möglichkeit Tools zu verwenden, die Seiten miteinander vergleichen und einen Ähnlichkeitsfaktor angeben ¹¹.

4.4 Massively Linking from Outside

Definition:

Es gibt viele verschiedene Methoden um die Popularität der eigenen Seite durch eine erhöhte Anzahl von eingehenden Links zu verbessern. Hierbei ist zu unterscheiden zwischen akzeptierten Verlinkungstechniken (zum Beispiel Verzeichnisse), unerwünschten und unerlaubten Techniken. Werden gewöhnliche Methoden zu sehr strapaziert, fallen sie ebenso in den Bereich der unerwünschten Techniken. In mäßigem Gebrauch akzeptierte Methoden sind bezahlte und gegenseitige Verlinkungen sowie Linkkreise und -verbünde. Das Erstellen von Verlinkungen auf die eigene Seite durch Gästebücher, Wikis und Communities wird als Spam gewertet, wenn Sie nicht themenrelevant sind. Nicht erwünschte Verknüpfungen sind *Link Farms* und kostenlose Linkseiten ¹². Böswillige und unerlaubte Methoden sind das Ausnutzen von schlecht erstellten Webseiten, bei denen sich Sicherheitslücken bilden. Mit *Cross-Site Scripting* und *HTML Injection* können auf diesen Seiten unerlaubterweise Verlinkungen in die HTML-Seite eingebaut werden.

Einfluss auf Suchmaschinen:

Die gängigsten Methoden - wie zum Beispiel *Link Farming* oder *Guestbook Spamming* - werden von den Suchmaschinen ohne Probleme erkannt [16]. Diese Links werden ignoriert, führen aber meistens zu keinem negativen Einfluss auf die Suchergebnisse. Das führt zu der Tatsache, dass es sich für Webseitenbetreiber nicht lohnt, die vorher genannten Link Kampagnen durchzuführen, da der Wert der Seite nicht gesteigert wird. Auf der anderen Seite lohnt es sich nicht für Mitstreiter durch diese Methoden ihre Konkurrenten auszuschalten, weil meistens kein negativer Einfluss bewirkt werden kann. Schwieriger ist es bei bezahlten Links und böswilligen Methoden durch Cross Site Scripting und HTML Injection. Diese Methoden können nur sehr schwer erkannt werden. Wenn diese Links nicht übermäßig auf einer Seite vorkommen und themenrelevant sind, sind sie nicht zu unterscheiden von gewöhnlichen Links.

Lösungsmöglichkeiten:

Generell sollten die unter der Definition genannten Methoden nicht angewendet werden oder nur in geringem, moderatem Maße. Gegenseitige Verlinkungen sind in Ordnung, wenn es nicht zu häufig vorkommt. Besonders auf Linkringe (eine Ausprägung von gegenseitigem Verlinken, siehe Bild 3) sollte verzichtet werden ¹³. Verzeichnisseiten werden akzeptiert, da sie themenrelevant sind. Die

¹¹ vgl. <http://www.webconfs.com/similar-page-checker.php> (Abruf: 26.3.11)

¹² vgl. http://www.webopedia.com/TERM/L/link_farming.html (Abruf: 26.3.11)

¹³ vgl. <http://www.seomarketingworld.com/seo-faq/cross-linking.php> (Abruf: 26.3.11)

Themenrelevanz spielt eine sehr bedeutende Rolle bei der Verlinkung. Relevante Seiten, die verlinkt werden, oder ein Link, bei dem der Linktext mit dem Inhalt der eigentlichen Seite übereinstimmt, haben eine höhere Gewichtung. Damit auf der eigenen Seite keine unerwünschten Verlinkungen zu anderen Seiten gewichtet werden, wird das Attribut *rel='nofollow'* innerhalb Links oder das Tag *'noindex'* auf der eigenen Seite gesetzt. Mit dem erstgenannten Verfahren können zum Beispiel alle Verlinkungen in Gästebucheinträgen aus dem Bewertungsverfahren ausgeschlossen werden (zum Beispiel für die Berechnung des Pageranks einer Seite).

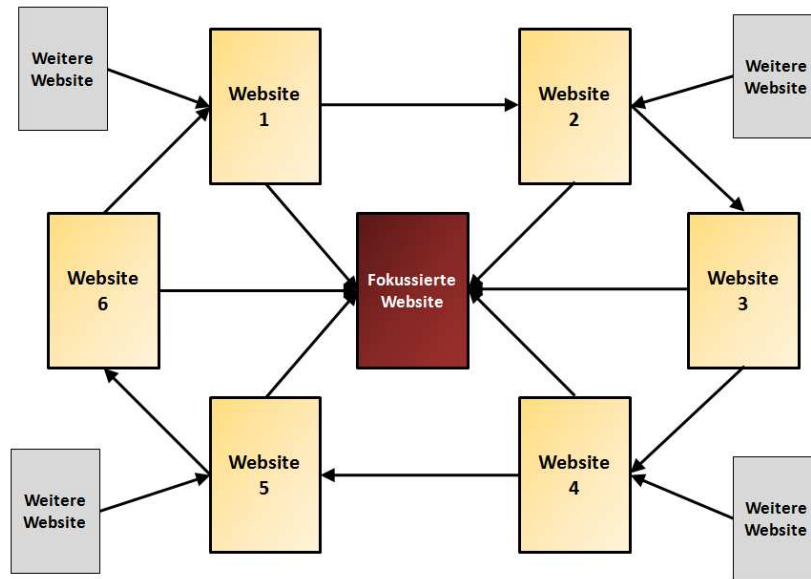


Abbildung 3. Beispielhafte Ausprägung eines Linkkreises

4.5 Poor Coding and Design

Definition:

Unter diesen Begriff fallen eine Anzahl an verschiedenen Methoden. Die Tabelle 2 zeigt die hierzu gefundenen Methoden. Die Vorgehen unterscheiden sich deutlich, weisen jedoch, wenn mehrere gleichzeitig angewendet werden, zusammen die größten negativen Einflüsse auf. Aus diesem Grund wurde dieser Begriff als Worst Practice an der Position 1 aufgeführt.

Als Paradebeispiel wird in diesem Fall das Problem der kanonischen Adresse beschrieben. Darunter versteht man im Bereich der Suchmaschinenoptimierung genau eine Adresse für eine Inhaltsseite auf einer Homepage zu besitzen [17].

Verschiedene Adressen können zum Beispiel unter folgenden Situationen auftreten: Bei der Benutzung von *Session IDs* im Adressfeld, bei der Benutzung von *Tracking Code* im Adressfeld oder wenn die Seite mit und ohne das Prefix *www* erreichbar ist.

Tabelle 2. Weitere Worst Practices im Bereich *Poor Coding and Design*

Zu viele Bilder, Videos und Podcasts	Falsche Meta-Tags
Session Ids	Fehlerseiten in Suchergebnissen
Sperren in robots.txt	Pop-Ups
Gleiche Titel Tags	Bearbeitungsseiten
Explodierende Adressen	Hosting-Auszeiten
Länderwahl als Startseite	Frames
Flash Homepages	Navigation mit Javascript oder Flash
Keine kanonischen Adressen	Unlesbare Adresse

Einfluss auf Suchmaschinen:

Sehr häufig werden schlecht programmierte Seiten nur teilweise oder gar nicht indexiert, weil die Suchmaschinen keinen Zugriff auf alle Seiten erhalten. Ferner ist es den Suchmaschinen nicht möglich, die Schlüsselbegriffe und entsprechenden Themen ausfindig zu machen. Dadurch treten die Seiten in den für sie relevanten Suchergebnissen nicht auf. Im schlimmsten Fall werden ohne Verbesserungen des Webseitenbetreibers für einen langen Zeitraum keine Ergebnisse der Seite angezeigt.

Bei inkonsistenten Adressen kann eine Suchmaschine nicht wissen, welche Seite indexiert werden soll. Dies wird nicht als Strafe angesehen, es entspricht nicht der Aufgabe eines Webseitenbetreibers einzigartige und relevante Ergebnisse zu liefern. Das Problem lässt sich mit Inhaltsduplikaten vergleichen, in diesem Fall treten die Duplikate jedoch aus Fahrlässigkeitsgründen auf. Ferner gibt es nur einen Inhalt, der über mehrere Adressen erreichbar ist. Es ergeben sich drei unterschiedliche Probleme [17]: Ist ein Artikel über verschiedene Adressen verfügbar, führt es dazu, dass Besucher diesen Artikel auf unterschiedlichen Adressen lesen. Wenn Sie einen Link zu diesem Artikel setzen möchten, verlinken sie folglich auf verschiedene Adressen. Dadurch wird für Suchmaschinen die Popularität des Artikels (in diesem Fall vereinfacht gemessen an der Anzahl eingehender Links von außen) in die Anzahl der verschiedenen Artikel gespalten. Weiterhin kann die Suchmaschine nicht entscheiden, welche der Adressen angezeigt werden soll. Haben beide Adressen einen ähnliche Popularität für die Suchmaschine, kann es sein, dass die Anzeigen der Adressen hin- und herwechseln. Dies kann dazu führen, dass diese Seiten nicht altern, was sich negativ auf die Positionierung in Suchergebnissen auswirken kann. Das dritte Problem entsteht

bei der Wartung der Homepage. Während des Verschiebens von Seiteninhalten müssen anstatt einer kanonischen Adresse alle Adressen gesucht werden, die auf diesen Artikel verweisen. Dies führt zu deutlichem Mehraufwand.

Lösungsmöglichkeiten:

Um *Poor Coding and Design* zu vermeiden, ist es ratsam sich *Best Practices* anzueignen und sich ausgiebig über Suchmaschinenoptimierung zu informieren. Das Auffinden von Fehlern lässt sich durch Suchmaschinentools beschleunigen. Ein Beispiel hierfür sind die *Google Webmaster Tools* ¹⁴.

Kanonische Adressen können ebenso unter den *Webmaster Tools* von Google überprüft werden. Eine weitere Möglichkeit ist das Betrachten von den Suchergebnissen der eigenen Website in Suchmaschinen. Bei letzterem ist es wichtig zu beachten, dass bereits Duplikate entfernt worden sein könnten. Taucht eine Seite unter mehreren Adressen auf, geht man von nicht-kanonischen Adressen aus [18]. Gleiche Titel sind ein Indiz hierfür. Die fehlerhaften Adressen können auf verschiedene Weisen korrigiert werden [19]:

Es ist möglich einer Seite unter einer bestimmten Adresse mittels einem *Meta-Tag* mitzuteilen, dass sie nicht indexiert werden soll. Wenn man möchte, dass alle Adressen mit der Variable *referrer* nicht indexiert werden sollen, kann dies der Suchmaschine über folgenden Quelltext mitgeteilt werden:

```
if (.$_GET['referrer']) {
echo "<meta name='robots' content='noindex,nofollow'>"
}
```

Ferner können bestimmte Seiten über die Datei *robots.txt* aus dem Index entfernt werden. Diese Datei sollte von jeder Suchmaschine gelesen werden, bevor sie eine Homepage durchsucht. Widerum Seiten mit der Variable *referrer* werden durch folgenden Quelltext entfernt:

```
User-agent: * Disallow: /*?referrer=
```

Als dritte Möglichkeit dienen Weiterleitungen. Hierzu muss der Seitenbetreiber einen Zugang zum Server besitzen. Es ist möglich mithilfe der Datei *.htaccess* sogenannte *301 Redirects* einzurichten. Es wird eine Weiterleitung von der unerwünschten auf die erwünschte Seite angegeben. Die Suchmaschinen leiten teilweise auch die Popularität der Seiten um:

```
RewriteCond %{HTTP_HOST} ^irgendeineseite.com
RewriteRule (.*?) http://www.irgendeineseite.com/$1 [R=301,L]
```

5 Schlusswort

5.1 Mangelnde Suchmaschinenoptimierung

Es hat sich herauskristallisiert, dass Methoden aus dem Bereich *Poor Coding and Design* eine große Rolle für die Anzeige der eigenen Seite innerhalb der Suchmaschinenergebnisse spielen. Meist werden diese Praktiken durch das Einhalten

¹⁴ vgl. <http://www.google.com/webmasters/tools/> (Abruf: 26.3.11)

von Best Practices vermieden. Für Personen, die keine direkte Suchmaschinenoptimierung anwenden, sind folgende drei Regeln wichtig: Zeige Suchmaschinen den gleichen Inhalt, der auch für Besucher gilt. Vermeide Methoden, welche unter Spamming und Blackhat fallen. Drittens lerne aus aufgetretenen Worst Practices, um diese in Zukunft zu vermeiden.

Wie bereits erwähnt, haben Suchmaschinen Probleme bei der Indexierung dieser Seiten. Zusätzlich sorgen Filter der Suchmaschinen dafür, dass schlecht optimierte Seiten eventuell nicht indexiert werden. Die Suchmaschinen gehen in diesem Fall von unabsichtlichen Methoden aus. Sie indexieren die entsprechende Seite nicht und strafen keine weiteren Seiten dieser Homepage ab. Häufig reicht es aus, die entsprechenden Fehler zu beheben. Bei der nächsten Indexierung der Homepage werden die korrigierten Seiten automatisch wieder in den Index aufgenommen.

5.2 Blackhat Methoden und Suchmaschinen-Spamming

Es gibt keine klar definierte Grenze zwischen nachhaltiger Suchmaschinenoptimierung und verbotenen Methoden. Die unterschiedliche Auslegung der Suchmaschinen verstärkt diese Tatsache. Um Konkurrenten keine Angriffsfläche zu bieten, ist es wichtig über diverse Blackhat und Spamming Methoden Bescheid zu wissen. Welche Maßnahmen als unzulässig betrachtet werden, kann bei den jeweiligen Suchmaschinen in Erfahrung gebracht werden. Es hat sich gezeigt, dass das Anwenden von Blackhat Methoden oder Suchmaschinen-Spamming zu keinem nachhaltigen Wettbewerbsvorteil führt. Die gängigen Suchmaschinen führen regelmäßig Updates durch, um neue unerlaubte Methoden zu unterbinden. Ein Erfolg ist demnach nur für eine kurze Zeit möglich. Wird erkannt, dass es sich um absichtliche Aktionen handelt, geht der Betreiber das Risiko einer Herabstufung in der Positionierung oder einer Sperre der kompletten Homepage ein [20]. Welche Herabstufungen auftreten können, werden im Folgenden kurz angedeutet [21]. Bei der Suchmaschine Google wurde beobachtet, dass Seitenergebnisse teilweise um 5 Stufen herabgesetzt wurden. Jedoch konnte hierfür kein driftiger Grund gefunden werden. Daher wird vermutet, dass diese Herabsetzung einem Update des Rankingverfahrens verschuldet ist [22]. Bei einer beobachteten Herabsetzung von 30 Positionen wird davon ausgegangen, dass auf diesen Seiten einige der in diesem Artikel genannten Praktiken bewusst angewendet worden sind (zum Beispiel *Guestbook Spamming* oder *Doorway Pages*). Schließlich wurde auch von Herabstufungen um 950 Positionen berichtet. Diese wurden nie offiziell bestätigt. Es wird davon ausgegangen, dass es sich hierbei um exzessives, unerlaubtes Optimieren sowie Blackhat Techniken handelt. Hier wird deutlich, dass nicht nur Blackhat Methoden, sondern auch zu starke Optimierungen in direkter Verbindung mit Bestrafungen für Seitenbetreiber stehen.

Während es bei leichteren Problemen ausreicht, diese zu beheben, muss bei erkannten Blackhat oder Spamming Methoden, welche meist zu einer dauerhaften Bestrafung führen, ein *Reinclusion Request* ausgeführt werden. Hierbei ist es wichtig offen darzulegen, dass die Ursache für die Bestrafung beseitigt wurde [23].

5.3 Angewandte Filter der Suchmaschinen

Folgende Filter sind unter Google bekannt und decken mangelnde Suchmaschinenoptimierung, Suchmaschinen-Spamming oder Blackhat Methoden auf [21]:

Tabelle 3. Bekannte Suchmaschinenfilter

Filter	Beschreibung
<i>Duplicated Content</i>	Erkennt Inhaltsduplikate.
<i>Bad Linking</i>	Erkennt nicht themenrelevante Links, Linkfarmen, Links auf ungläubige Seiten etc...
<i>Cross Linking Penalty</i>	Erkennt mehrere gegenseitig verlinkte Seiten.
<i>Excessive Use of Same Anchor Text</i>	Erkennt gleiche Ankertexte bei Verlinkungen.
<i>Excessive Use of Reciprocal Linking</i>	Erkennt beidseitige Verlinkungen.
<i>Thin Affiliate Site</i>	Erkennt Seiten, die sehr viele Affiliate Links und geringfügigen Inhalt aufweisen.
<i>Keyword Stuffing</i>	Erkennt das Überladen von Seiten mit Schlüsselbegriffen.
<i>Hidden Text or Links Penalty</i>	Erkennt versteckte Texte und Links.
<i>Automated Page Redirects</i>	Erkennt automatisierte Weiterleitungen, die nicht vom Besucher ausgelöst werden.

5.4 Weitere Themengebiete

Innerhalb dieser Arbeit wurden Spamming-, Blackhat- sowie unbeabsichtigte Methoden beschrieben. Das Thema wurde zu *Enterprise Search* sowie *mobile SEO* abgegrenzt.

Als erweitertes Themengebiet kann die Kombination von Enterprise Search und SEO Worst Practices angesehen werden. Die Unterschiede von Enterprise Search und SEO wurden in dem Unterkapitel 1.2 bereits erläutert. Daraus ergeben sich folgende Fragestellungen: Ist es möglich SEO Praktiken, insbesondere Worst Practices, für Internetsuchmaschinen auf Unternehmenssuchmaschinen abzubilden? Besteht ein Mehrwert für die einzelnen Mitarbeiter und das gesamte Unternehmen? Können und sollen Dokumente innerhalb eines Unternehmens für das Internet verfügbar gemacht werden? Falls ja, inwiefern sollten SEO Praktiken angewendet werden und bringen diese einen erheblichen Mehrwert?

Ein weiteres Themenfeld ist *mobile SEO*. Im Unterkapitel 1.2 wurden bereits die Motive hierfür genannt. Auf diesem Thema wird in Zukunft ein erhöhter Fokus liegen.

Literatur

1. Schuster, J.: Google sperrt nun auch deutsche webseiten mit versteckten suchwörtern aus <http://heise.de/-172419> (Abruf: 25.03.2011).
2. Hitzelberger, F.: Sabotage - manipulation durch google bowling /url<http://www.domain-recht.de/magazin/domain-news-2005/sabotage-manipulation-durch-google-bowling-id522.html> (Abruf: 25.03.2011).
3. Segal, D.: A bully finds a pulpit on the web http://www.nytimes.com/2010/11/28/business/28borker.html?_r=1&pagewanted=all (Abruf: 25.03.2011).
4. Singhal, A.: Being bad to your customers is bad for business <http://googleblog.blogspot.com/2010/12/being-bad-to-your-customers-is-bad-for.html> (Abruf: 25.03.2011).
5. Jaimie Sirovich, C.D.: Professional Search Engine Optimization with PHP. Wiley Publishing, Inc. (2007)
6. Yvonne von Bischoffinck, M.C.: Suchmaschinen-Marketing: Konzepte, Umsetzung und Controlling. Springer (2007)
7. Nikolaus Mohr, A.K.S.B.: Mobile web watch-studie 2010 - durchbruch auf raten: Mobiles internet im deutschsprachigen raum (2010) http://www.accenture.com/SiteCollectionDocuments/Local_Germany/PDF/Accenture_Mobile_Web_Watch_2010.pdf (Abruf: 25.03.2011).
8. Wirtschaftslexikon, G.: Gabler wirtschaftslexikon, stichwort: Best practice <http://wirtschaftslexikon.gabler.de/Archiv/122484/best-practice-v3.html> (Abruf: 25.03.2011).
9. Erdmeier, P.: Lexikon suchmaschinenoptimierung: Black hat seo <http://www.lexikon-suchmaschinenoptimierung.de/black-hat-seo.htm> (Abruf: 25.03.2011).
10. Bahrs, J.: Enterprise Search - Suchmaschinen für Inhalte im Unternehmen. In Handbuch Internet-Suchmaschinen: Nutzerorientierung in Wissenschaft und Praxis. Akademische Verlagsgesellschaft AKA (2009)
11. Rajat Mukherjee, J.M.: Enterprise search: Tough stuff. Queue **2** (April 2004)
12. Kriest, R.: Keyword-stuffing: Wie man den keyword-overkill in postings vermeidet <http://www.prologgerworld.de/keyword-stuffing-vermeiden-live-keyword-analysis/439/> (Abruf: 25.03.2011).
13. Malaga, R.A.: Worst practices in search engine optimization. Communications of the ACM **51** (2008)
14. Patalong, F.: Suchmaschinen: Wie sich bmw ausgoogelte <http://www.spiegel.de/netzwelt/web/0,1518,399267,00.html> (Abruf: 25.03.2011).
15. Lasnik, A.: Duplicate content - best practice bei dupliziertem inhalt http://googlewebmastercentral-de.blogspot.com/2007/03/duplicate-content-best-practice-bei_08.html (Abruf: 25.03.2011).
16. Ott, S.: Link farms und guestbook links - suchmaschinenmanipulation <http://www.linksandlaw.de/suchmaschinen-leitfaden-zur-Manipulation-von-Suchergebnissen-linkfarms.htm> (Abruf: 25.03.2011).
17. Lurie, I.: Canonicalization defined (part 1 of 3) — internet marketing strategy: Conversation marketing <http://www.conversationmarketing.com/2009/09/seo-101-canonicalization-1.htm> (Abruf: 25. 03. 2011).
18. Lurie, I.: Detect canonicalization problems (part 2 of 3) — internet marketing strategy: Conversation marketing <http://www.conversationmarketing.com/2009/09/seo-101-canonicalization-part-2.htm> (Abruf: 25. 03. 2001).

19. Lurie, I.: Fix canonicalization problems (part 3 of 3) — internet marketing strategy: Conversation marketing <http://www.conversationmarketing.com/2009/10/how-to-fix-canonicalization.htm> (Abruf: 25.03.2011).
20. Ott, S.: Leitfaden zur manipulation von suchmaschinen <http://www.linksandlaw.de/suchmaschinen-leitfaden-zur-Manipulation-von-Suchergebnissen.htm> (Abruf: 25.03.2011).
21. Dirks, B.: Google search engine ranking penalties <http://www.seonightmares.com/googlepenalties.htm> (Abruf: 25.03.2011).
22. WALL, A.: Google ranking -6 penalty filter <http://www.seobook.com/google-ranking-6-penalty-filter> (Abruf: 25.03.2011).
23. Erdmeier, P.: Lexikon suchmaschinenoptimierung: Reinclusion request <http://www.lexikon-suchmaschinenoptimierung.de/reinclusion-request.htm> (Abruf: 25.03.2011).

6 Anhang

Tabelle 4: Gefundene Worst Practices

Pos.	Worst Practice	Quellen-Ids	Gewichtung
1	Cloaking	2,3,4,5,6,11,12,13	9
2	Doorway Pages	2,3,4,11,12,13	7
3	Invisible Elements	2,3,4,5,6,11,13	7
4	Keyword Stuffing	2,3,4,5,6,12,13	7
5	Link Farms	6,11,12,13	5
6	Duplicate Content	2,3,4,6,13	5
7	Paid Links to Pass Pagerank	3,5,6,8	5
8	Pagejacking/Hijacking	4,6,12,13	5
9	Content Generators	4,5,6,11	4
10	Wrong Redirects	2,3,4,5	4
11	Content Separation	2,10	3
12	Guestbook Spamming	11,12,13	3
13	Bowling	11,13	3
14	Cross Linking	2,3,5	3
15	Single Pixel links	2,3,13	3
16	Using Frames	2,4,5	3
17	Dynamic, Long & Unreadable URLs	2,3,4	3
18	Link Spamming (Wiki, Blog, Community)	2,13	2
19	Toilet Bowl of Death (Exploding URLs, Spider Trap)	1	2
20	Canonicalization	7	2
21	Keyword Dilution	2,3	2
22	Many Outgoing Links	2,3	2
23	Hidden Links	3,13	2
24	Illegal Content	2,3	2
25	Session IDs	2,4	2
26	Bans in robots.txt	2,3	2
27	Same Title Tags	4,13	2

Fortsetzung auf der nächsten Seite

Tabelle 4: Gefundene Worst Practices

Pos.	Worst Practice	Quellen-Ids	Gewichtung
28	Submit Site Repeatingly to Search Engines	4,6	2
29	Free For All Linking Pages	4,5	2
30	Article Spinning	14	2
31	Flash Homepage	1,4	2
32	Competitors Brand Names in Meta-Tags	4,6	2
33	PageRank as Primary Metric for SEO	9	2
34	Google Bomb	12	2
35	Doorway Domains	12	2
36	Referrer Log Spamming	13	1
37	Blog-Ping	11	1
38	Meta Refresh	2	1
39	Outbound Links to Farms	2	1
40	HTML Injection (e.g. Web-Glimpse)	11	1
41	Poor Coding and Design	2	1
42	Too much Images, Podcasts and Videos	2	1
43	Flash Movies	2	1
44	Hosting on a blacklisted IP address	2	1
45	Hosting Downtown, 404 Error Pages, Maintenance sites	2	1
46	Blocking Internet Archive	3	1
47	Bad Neighbors	3	1
48	Navigations with Pull Down Boxes, Flash or Javascript	4	1
49	Splash page	4	1
50	Multiple Domains	4	1
51	Pop-ups	4	1
52	Click Here Links	4	1

Fortsetzung auf der nächsten Seite

Tabelle 4: Gefundene Worst Practices

Pos.	Worst Practice	Quellen-Ids	Gewichtung
53	Superfluous Text like ‘Welcome To’	4	1
54	Irrelevant Keywords	4	1
55	Spamglish	4	1
56	Buying Expired Domains	4	1
57	Country Selector as Your Homepage	4	1
58	Cross Site Scripting	5	1
59	Scraper Sites	13	1
60	Image Links	1	1
61	Repetition of Keywords in Internal Anchor Text	3	1
62	Error Pages in Search Results	4	1
63	File Not Found Error Returns a 200 Status Code	4	1
64	Reciprocal Link Trading	5	1
65	Bottom Feeder Links	5	1
66	Page Dedication to Unusable Keywords	6	1
67	Wrong Meta-Tags	13	1
68	Sybill Attack	13	1

Tabelle 5: Verwendete Literatur für das Finden von Worst Practices

Id	Herkunft	Gewichtung
1	http://www.conversationmarketing.com/2010/08/seo-worst-practices-duplication-death.htm (Abruf: 25.03.2011)	2
2	http://www.webconfs.com/15-minute-seo.php (Abruf: 25.03.2011)	1
3	http://www.seonightmares.com/badseopractice.htm (Abruf: 25.03.2011)	1
4	http://searchengineland.com/seo-checklist-of-best-and-worst-practices-part1-43752 (Abruf: 25.03.2011)	1
5	http://www.seotier.com/23/seo-worst-practice-manual/ (Abruf: 25.03.2011)	1
6	http://www.dontdrinkthekoolaidblog.com/search-engine-optimization-worst-practices/ (Abruf: 25.03.2011)	1
7	http://www.conversationmarketing.com/2009/09/seo-101-canonicalization-1.htm (Abruf: 25.03.2011)	2
8	http://www.google.com/support/webmasters/bin/answer.py?hl=en&answer=66736 (Abruf: 25.03.2011)	2
9	http://www.seomoz.org/blog/the-science-of-ranking-correlations (Abruf: 25.03.2011)	2
10	http://googlesystem.blogspot.com/2006/08/content-separation.html (Abruf: 25.03.2011)	2
11	Ross A. Malaga, Worst Practices in Search Engine Optimization, Communications of the ACM, 2008	2
12	http://www.linksandlaw.de/suchmaschinen-leitfaden-zur-Manipulation-von-Suchergebnissen.htm (Abruf: 25.03.2011)	2
13	http://de.wikipedia.org/wiki/Suchmaschinen-Spamming (Abruf: 25.03.2011)	1
14	http://www.myseosolution.de/seo-tools/article-wizard/#article-spinner (Abruf: 25.03.2011)	2

Tabelle 6: Kategorisierungsmöglichkeit der Worst Practices

Kategorie	Methoden
Cloaking	Content Separation, Doorway Pages, Meta Refresh
Content Related	Duplicate Content, Illegal Content, Content Generators, Invisible Content, Too much Images/Videos/Podcasts, Superfluous Text, Splash Pages, Spamglisch, Flash Movies, Page Dedication to unusable Keywords
Index Related	Blog-Ping, Multiple Domains, Repeating Site Submit, Page Hijacking, Buying of Expired Domains, Spamming of Referrer Logs, Scraper Sites, Article Spinning
Keyword Related	Irrelevant Keywords, Keyword Dilution, Keyword Stuffing, Invisible Keywords, Repetition of Keywords in Internal Anchor Text
Link Related	Hidden Links, Single Pixel Links, Cross Links, Too many Outgoing Links, Link Spamming, Link Farms, Comment Spamming, Wiki Spamming, Paid Links, Outbound Links to Farms, Guestbook Spamming, Click Here Links, Free For All Linking Pages, Image Links, Reciprocal Link Trading, Bottom Feeder Links, Google Bombing
Poor Coding and Design	Use of Frames, Use of Flash, Use of Popups, Country Selector as Home Page, Cross Site Scripting Vulnerability, Navigations Consisting of Pull Down Boxes/Flash/JavaScript, Same Title Tags, No Meta Tags, HTML Injection, Error Pages in Search Results, Incorrect Status Codes, Wrong Meta Tags
URL Related	Use of Session Ids in URLs, Wrong Redirects, Dynamic/Long/Unreadable URLs, Canonicalization, Exploding URLs
Others	Bad Neighbours, Bans in robots.txt, Blocking Internet Archive, Hosting Downtowns or Maintenance Sites, Hosting on Blacklisted IP address, Competitors brand names in Meta Tags