

Diplomarbeit

Suchmaschinen-Marketing –
Methoden, um in Suchmaschinen
gefunden zu werden.

von

Hans-Peter Bauer

betreut von

Dipl.Soz./Wiss.Dok. Jutta Bertram

im Fachbereich Information- und Knowledge Management

Fachhochschul-Studiengang Informationsberufe
Eisenstadt 2006

Eidesstattliche Erklärung

Ich habe diese Diplomarbeit selbstständig verfasst, alle meine Quellen und Hilfsmittel angegeben, keine unerlaubten Hilfen eingesetzt und die Arbeit bisher in keiner Form als Prüfungsarbeit vorgelegt.

Walbersdorf, im Mai 2006

Kurzreferat

Viele Studien unterstreichen die Bedeutung von algorithmischen Suchmaschinen für die private und geschäftliche Informationssuche im Internet. Deshalb ist es für viele Webseiten von besonderer Wichtigkeit, auf der ersten Ergebnisliste der bedeutendsten Suchmaschinen vertreten zu sein. Ziel der vorliegenden Arbeit ist es, diese Entwicklung anzusprechen, indem die Möglichkeiten von Suchmaschinen-Marketing vorgestellt werden. Neben dem Instrument der keywordbezogenen Werbung, wird vor allem die Suchmaschinen-Optimierung diskutiert und analysiert. Insbesondere die Hypothese, dass Webseiten, die in den Ergebnislisten von Suchmaschinen besonders gut platziert sind, sich sowohl durch ausgeprägte Anwendung von On-The-Page Methoden, als auch durch inhaltliche Qualität auszeichnen, wird geprüft.

Details zu den zwei Instrumenten von Suchmaschinen-Marketing, nämlich Suchmaschinen-Optimierung und keywordbezogene Werbung, werden mit Hilfe von Fachliteratur erarbeitet. Dazu werden relevante Themen des Suchmaschinen-Marketing, basierend auf technischer, wissenschaftlicher und wirtschaftlicher Information, hervorgehoben. Um die Hypothese zu überprüfen werden HTML-Quellcodes von Webseiten, die im Rahmen eines Suchmaschinen-Optimierungswettbewerbs besonders gut platziert waren, qualitativ untersucht. Somit kann auf jene Form der Optimierung geschlossen werden, die zu langfristigen Erfolg bei den drei Suchmaschinen Google, Yahoo! und MSN führt.

Die Ergebnisse zeigen, wie es möglich ist, eine Website in allen drei bedeutenden Suchmaschinen sogar in einem stark umkämpften Wettbewerbsumfeld zu platzieren. Die Ergebnisse zeigen wichtige Faktoren der Optimierung auf und ermöglichen Rückschlüsse auf die Rolle der Suchmaschinen-Optimierung und der Qualität des Inhalts für den Erfolg einer Webseite. Für weitere Untersuchungen wird empfohlen, die gefundenen Methoden außerhalb eines Wettbewerbs anzuwenden und ihre Richtigkeit zu testen.

Schlagwörter: Suchmaschine, Internet-Marketing, Suchmaschinen-Marketing, Keywordbezogenes Marketing, Suchmaschinen-Optimierung

Abstract

According to various studies, crawler-based search engines are the most important retrieval tools on the Internet today. Due to this, hardly any website can afford not being listed in the result pages of the major search engines. The aim of this diploma thesis addresses this development by presenting the possibilities offered by search engine marketing to promote a website. Apart from the option of keyword-related advertising, search engine optimisation is discussed and analysed. Especially the hypothesis that well-ranked webpages depend on both, the use of on-the-page methods of search engine optimisation and good content, is tested.

A review of current literature on this topic reveals details about search engine optimisation and keyword advertising, the two applications of search engine marketing. Furthermore, based on technical, scientific and economical information about search engines, main issues of search engine marketing are highlighted. In order to discuss the given hypothesis, HTML source codes of well-ranked websites are analysed qualitatively to investigate, which factors lead to long-term success with each of the three search engines Google, Yahoo and MSN.

The results of the analysis indicate how it is possible to rank a website in all major search engines, even in a very competitive environment. Though the most important factors of optimisation are suggested, the meaning of quality and uniqueness of the content are emphasised, as well. For further research on this topic, it is recommended to test the methods by applying them in a less competitive surrounding.

Keywords: search engine, Internet marketing, search engine marketing, keyword advertising, search engine optimisation

Executive Summary

Die vorliegende Arbeit verfolgt das Ziel, dem Leser die Möglichkeiten vorzustellen, die Suchmaschinen-Marketing bietet, um eine Webseite über die Ergebnislisten der Suchmaschinen zu bewerben. Vor dem Hintergrund dieser zentralen Forschungsfrage wird anhand einer Literaturrecherche das technische, informationswissenschaftliche und wirtschaftliche Umfeld von Suchmaschinen erläutert. Im Mittelpunkt der Erörterungen stehen hier der technische Aufbau von Suchmaschinen und deren Aufgaben, die angewandten Prinzipien des Information Retrieval sowie die Entwicklung des Suchmaschinenmarktes bis heute. Anschließend werden keywordbezogene Werbung und Suchmaschinen-Optimierung als Instrumente des Suchmaschinen-Marketing vorgestellt. Wiederum unter Bezugnahme auf facheinschlägige Autoren, wird detailliert und anschaulich auf wichtige Aspekte der keywordbezogenen Werbung und kritische Faktoren der Suchmaschinen-Optimierung eingegangen.

In einer empirischen Untersuchung, in der die Quelltexte von elf besonders erfolgreichen Webseiten in Bezug auf inhaltliche Besonderheiten und angewandte Methoden der Suchmaschinen-Optimierung qualitativ untersucht werden, wird die Hypothese „Webseiten, die in den Ergebnislisten Suchmaschinen besonders gut platziert sind, zeichnen sich sowohl durch ausgeprägte Anwendung von On-The-Page Methoden, als auch durch inhaltliche Qualität aus.“ geprüft. Dabei liegt der Schwerpunkt der Untersuchung auf Faktoren, die aus dem Quelltext ersichtlich sind, also auf On-The-Page Kriterien. Hingegen müssen die Angaben zu den Off-The-Page Kriterien, die die zweite Säule der Suchmaschinen-Optimierung bilden, aus Mangel an verlässlichen Daten mit einer gewissen Vorsicht betrachtet werden.

Im speziellen Fall der elf untersuchten Webseiten konnten vier erfolgreiche Typen von Webseiten abgeleitet werden. Sie weisen jeweils eine unterschiedliche Intensität der Optimierung und inhaltliche Schwerpunkte auf. Dazu zählen unter anderem Webseiten, die nicht optimiert sind und sehr ansprechende und seriöse Inhalte anbieten, aber auch eine Webseite, die keinen sinnvollen Text enthält und nur durch Optimierungsmaßnahmen in die oberen Bereiche der Suchmaschinenergebnislisten gelangt ist. Die Hypothese ist daher zu falsifizieren. Die Ergebnisse der Untersuchung lassen also einerseits Rückschlüsse auf die Rolle der Suchmaschinen-Optimierung zu, andererseits zeigen sie anhand von

praktischen Beispielen, welche On-The-Page Methoden der Suchmaschinen-Optimierung potentiell erfolgsversprechend sind.

Zusammengefasst erfährt der Leser Informationen über die Funktion von Suchmaschinen und die Methoden, wie sich die Ergebnislisten der Suchmaschinen für die verbesserte Vermarktung einer Webseite nutzen lassen. Die qualitative Analyse vervollständigt das Bild von Suchmaschinen-Optimierung, indem es dem Leser erfolgreiche Maßnahmen veranschaulicht und die Rolle der Suchmaschinen-Optimierung auf Basis der Ergebnisse verdeutlicht.

Inhaltsverzeichnis

1	Einleitung	1
2	Internet-Suchmaschinen	3
2.1	Die Technik von Suchmaschinen	4
2.2	Suchmaschinen aus der Sicht der Informationswissenschaften . .	6
2.3	Die Entwicklung des Suchmaschinenmarkts	9
3	Suchmaschinen-Marketing	12
3.1	Definition und Ziel von Suchmaschinen-Marketing	13
3.2	Suchmaschinen-Werbung	16
3.2.1	Pay-Per-Click Suchmaschinen	16
3.2.2	Auswahl der Keywords	17
3.2.3	Formulierung von Anzeigentext und Landing Page	20
3.2.4	Gefahren und Risiken beim Einsatz von Paid Placements	22
3.3	Suchmaschinen-Optimierung	23
3.3.1	On-The-Page Methoden	24
3.3.2	Off-The-Page Methoden	26
3.3.3	Grenzen der Optimierung	28
3.4	Überwachung und Erfolgskontrolle von Suchmaschinen- Marketing-Kampagnen	32
4	Empirische Untersuchung von HTML-Quelltexten	34
4.1	Untersuchungsgegenstand und Ausgangssituation	35
4.2	Vorgehensweise	36
4.3	Ergebnisse der qualitativen Analyse	40
4.4	Zusammenfassung	53
4.4.1	On-The-Page Kriterien	54
4.4.2	Off-The-Page Kriterien	56
4.4.3	Intensität der Optimierung und inhaltliche Qualität . . .	59
4.5	Schlussfolgerungen	62
	Literatur	68

Abbildungsverzeichnis

1	Komponenten einer Suchmaschine am Beispiel von FAST.	5
2	Der Suchmaschinenmarkt in den Jahren 2002 und 2005.	10
3	Die Vernetzung der Suchmaschinen in Deutschland.	11
4	Die Vernetzung der Suchmaschinen weltweit.	11
5	Klassische Ergebnisse und bezahlte Einträge am Beispiel der Ergebnisseite der Suchmaschine Yahoo!	14
6	Die Beziehung von Keywords, Textanzeige und Landing Page. .	20
7	Grundsätzlicher Aufbau einer HTML-Datei.	24
8	Beispiel für die unzulässige Methode des Keyword-Stuffing. . . .	42
9	Beispiel für die Positionierung von Keywords im Kopf eines HTML-Dokuments.	43
10	Beispiel für die Möglichkeiten, wie Keywords in Links und Bil- dern untergebracht werden können.	44
11	Blindtext und versteckte Links auf den Seiten von Daniel Gal. .	46
12	Beispiel für die Möglichkeit Keywords in einem Kommentar zu platzieren.	48
13	Ausschluss eines Robots durch das Robots Exclusion Protocol in einer robots.txt-Datei.	49
14	Ausführliche Erschließung eines Internet-Dokuments im Head einer HTML-Datei.	51
15	Beispiel für einen optimierten Text.	51
16	Grafische Darstellung der Webseiten in Bezug auf inhaltliche Beschreibung und Intensität der Suchmaschinen-Optimierung. .	64

Tabellenverzeichnis

1	Auswahl von PPC-Anbietern, die in Europa tätig sind.	17
2	Platzierungen der ausgewählten Websites in den betrachteten Suchmaschinen.	38
3	Zusammenfassung der Ergebnisse der qualitativen Untersuchung.	60

1 Einleitung

Obwohl das Internet viele Möglichkeiten bietet, Informationen zu finden, gelten Suchmaschinen als der wichtigste Einstiegspunkt für Internetrecherchen. Sowohl im privaten, als auch im geschäftlichen Bereich hat diese einfache und schnelle Möglichkeit, Informationen und Produkte zu finden, natürlich einen besonderen Reiz, wahrscheinlich auch darum, weil die Suchmaschine die riesige Menge an verfügbaren Dokumenten in eine übersichtliche Reihung bringt und damit dem Benutzer die Auswahl erleichtert.

Für den Betreiber einer Internetseite, der mit seinem Informationsangebot oder seinem Webshop gefunden werden will, entsteht aus dieser Entwicklung eine ernstzunehmende Herausforderung, denn nur die ersten Ergebnisseiten der Suchmaschinen erregen die Aufmerksamkeit der Suchmaschinenbenutzer und werden angeklickt. Die aussichtsreichsten Positionen sind also die ersten zehn, bestenfalls die ersten zwanzig Ergebnisse. Hier ist aber bei vielen Suchbegriffen der Wettbewerb groß und der Platz eng geworden, sodass ohne das Zutun des Webmasters kaum eine Seite mehr die Spitze der Ergebnislisten erreicht. Besonders, wenn das Gefundenwerden über den Erfolg eines Unternehmens entscheidet, wird man sich überlegen müssen, welche Mittel und Wege es gibt, um potentielle Kunden über die Suchmaschinenergebnisse auf die eigene Webseite zu bekommen.

In der vorliegenden Arbeit werden die Möglichkeiten keywordbezogene Werbung und Suchmaschinen-Optimierung, die sich hinter dem Titel Suchmaschinen-Marketing verbergen, vorgestellt. Ausgehend von technischen und informationswissenschaftlichen Grundlagen sowie der Darstellung der Situation auf dem Suchmaschinenmarkt, werden die allgemeinen Wesenszüge von Suchmaschinen-Marketing vorgestellt und insbesondere auf wichtige Faktoren bei der Anwendung der beiden Instrumente – keywordbezogene Werbung und Suchmaschinen-Optimierung – eingegangen. Basierend auf facheinschlägiger Literatur wird dem Leser ein umfassendes Bild der Thematik präsentiert.

Der empirische Teil der Arbeit (Kapitel 4) widmet sich schließlich der Suchmaschinen-Optimierung im Speziellen. Hier werden elf Quelltexte, die im Rahmen eines Suchmaschinen-Optimierungswettbewerbs besonders erfolgreich waren, qualitativ analysiert. Ziel der Untersuchung ist es einerseits, die theoretischen Angaben aus dem Kapitel 3.3 praktisch zu überprüfen und mit Beispielen

len zu veranschaulichen. Andererseits dienen die Ergebnisse dazu, die Rolle der Suchmaschinen-Optimierung im Verhältnis zur Qualität des Inhalts bei der Erstellung der Webseite einschätzen zu können. Insbesondere soll die Hypothese „Webseiten, die in den Ergebnislisten Suchmaschinen besonders gut platziert sind, zeichnen sich sowohl durch ausgeprägte Anwendung von On-The-Page Methoden, als auch durch inhaltliche Qualität aus.“ geprüft werden.

2 Internet-Suchmaschinen

Lewandowski (2005, 24) stellt fest, dass alle Suchwerkzeuge im World Wide Web entweder Ausprägungen von manuell erstellten Verzeichnissen oder von algorithmischen Suchmaschinen sind. Während also Verzeichnisse, wie etwa das Open Directory Project (ODP), von Menschen gepflegt werden und damit die Qualität der Einträge gut kontrolliert werden kann, erfolgt die Qualitätsbewertung bei algorithmischen Suchmaschinen, wie zum Beispiel die Websuche bei Google, durch einen Rechner, der hierfür unterschiedliche Kriterien heranzieht.

Auch wenn der Trend zu sogenannten hybriden Suchdiensten¹ und Portalen² geht, bezieht sich die vorliegende Arbeit ausschließlich auf die algorithmischen Suchmaschinen und legt den Fokus, etwa im Falle von Yahoo!, auf die Suchmaschine und nicht auf das Verzeichnis oder sonstige Dienste. Es steht also das Wesen von algorithmischen Suchmaschinen im Mittelpunkt der folgenden Ausführungen. Dabei wird in den kommenden Abschnitten die Suchmaschine aus verschiedenen Perspektiven beschrieben, um dem Leser ein möglichst umfassendes Bild über Suchmaschinen im Internet zu vermitteln. Während in den Publikationen, die als Quellen für diese Arbeit dienen, häufig nur eine Sicht auf Internet-Suchmaschinen dargestellt wird, beschreibt diese Arbeit das Wesen der Suchmaschine aus den drei Blickwinkeln.

Zu Beginn werden die **Suchmaschinen aus technischer Sicht** beschrieben, um einen Einblick in ihre Arbeitsweise zu bekommen. Anhand der Beschreibung der Komponenten und deren Aufgabe bekommt der Leser eine Einführung in die technischen Grundlagen und einen Eindruck vom Systemumfeld einer Suchmaschine.

Im Abschnitt **Suchmaschinen aus Sicht der Informationswissenschaften** werden jene Herausforderungen zusammengefasst dargestellt, mit denen Suchmaschinen als Information Retrieval Systeme konfrontiert sind. Darüber

¹ „Trend zu sog. hybriden Suchdiensten: Suchmaschinen bieten fast immer auch einen Katalog - und Kataloge leiten immer häufiger Suchanfragen auch zu Suchmaschinen weiter“ (Meier, 2006)

² „Trend zu sog. Portalseiten: Suchdienste bieten neben der reinen Suchfunktion weitere Dienstleistungen: E-mail, Tarifübersicht, Wettervorhersage, Nachrichten, Routenplanung etc.“ (Meier, 2006)

hinaus werden die informationswissenschaftlichen Prinzipien besprochen, die die Grundlage für das Funktionieren einer Suchmaschine bilden.

Schließlich wird auch auf die **Entwicklung des Suchmaschinenmarkts** eingegangen. Ausgehend von einem Rückblick auf die Ursprünge der heutigen Suchmaschinen, wird die heutige Situation des Suchmaschinenmarkts erläutert und insbesondere die wirtschaftliche Verflechtung der verschiedenen Suchdienste thematisiert.

2.1 Die Technik von Suchmaschinen

Auch wenn die Internet-Suchmaschinen, ebenso wie das Internet im Gesamten, ihre Wurzeln im universitären Umfeld haben (siehe Kapitel 2.3), werden die neuesten Erkenntnisse, die die Funktion und den Aufbau von Suchmaschinen betreffen, zum Großteil von den Anbietern der Suchmaschinenteknologie in den USA erarbeitet. Es ergibt sich also, wie auch Lewandowski (2005, 14) berichtet, folgendes Problem: Die Forschungsergebnisse liegen, sofern sie überhaupt öffentlich zugänglich sind, vor allem in englischer Sprache, in Form von Patenten oder fertig gestellten Anwendungen vor.

Auch wenn manche interessanten Details von den Suchmaschinenbetreibern unter Verschluss gehalten werden, kann über den grundsätzlichen Aufbau der Suchmaschinen, der sich unter den Anbietern kaum unterscheidet, gut Auskunft gegeben werden. So zeigt Abbildung 1 den kompletten Aufbau einer Suchmaschine auf Basis der Technologie des norwegischen Suchmaschinenanbieters FAST-Fast Search & Transfer. Diese dargestellten Module fasst Glöggler (2003, 25 ff.) grob in drei Komponenten zusammen, aus denen jede Suchmaschine besteht; und zwar das Webrobot-System, das Information Retrieval System und den Query Processor.

Das **Webrobot-System** ist „*diejenige Systemkomponente der Suchmaschinen, die für die Erfassung von neuen und veränderten Ressourcen im Internet verantwortlich ist*“ (Glöggler, 2003, 26). Indem ein automatischer Web Browser (Gatherer) ausgehend von einer Seite allen Links folgt, sammelt er Dokumente im World Wide Web und speichert deren Pfad in einer URL-Datenbank. Um eine effiziente Arbeitsweise sicherzustellen, organisiert ein weiteres Modul (Loader) die auszuführenden Aufträge. Als letztes Modul filtert der Checker,

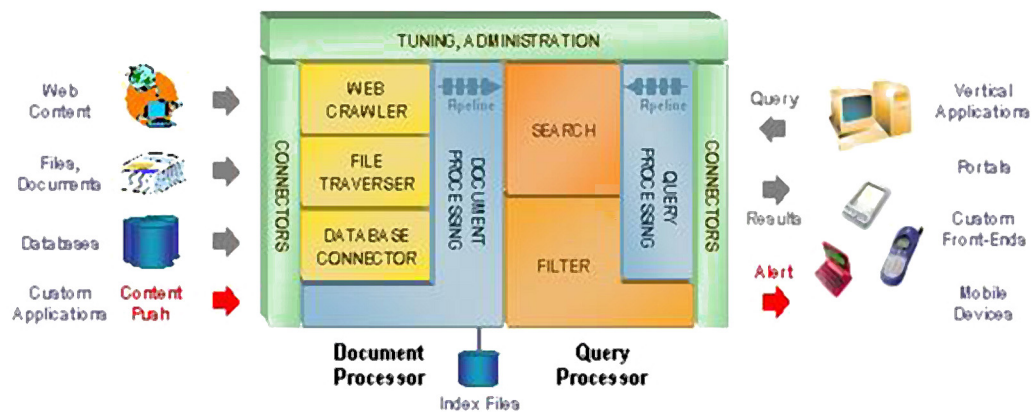


Abbildung 1: Komponenten einer Suchmaschine am Beispiel von FAST, vorgestellt von Aaberg (2006).

je nach Konfiguration der Suchmaschine, die gefundenen Informationen und sorgt vor allem dafür, dass keine unbrauchbaren Dokumente³ an das Information Retrieval System weitergegeben werden.

Das **Information Retrieval System** einer Suchmaschine hat die Aufgabe, die vom Webrobot-System gesammelten und übergebenen Daten zu normalisieren, zu analysieren und zu indexieren (Glöggler, 2003, 40). Bei der Normalisierung der Dokumente wird das Datenformat vereinheitlicht und in indexierbare Einheiten zerlegt. Als solche gelten zum Beispiel ganze Wörter, Wortstämme oder sogenannte N-Gramme⁴ (Lewandowski, 2005, 27). Schließlich werden die Deskriptoren gemeinsam mit ihrem Vorkommen im Dokument im Rahmen der Indexierung gespeichert. Es entsteht ein invertiertes Dateisystem aus den direkten Dateien, den indirekten Dateien und dem Index (Glöggler, 2003, 58). Während in den direkten Dateien jeweils ein Dokument mit allen vorkommenden Wörtern gespeichert wird, gibt es für jedes indexierte Wort eine invertierte Datei mit einer Referenz, in welchen direkten Dateien dieses Wort vorkommt. Der Index ist schließlich die alphabetische Liste aller Wörter mit einem Verweis auf die jeweilige invertierte Datei.

³ Auf der technischen Ebene des Webrobot-Systems werden hier vor allem Dokumente gelöscht, deren Dokumententyp oder Dateiformat vom IR-System nicht indexiert werden kann, die bereits indexiert worden sind (Dublettenerkennung) oder deren URL syntaktisch nicht richtig ist.

⁴ Anstelle der Worterkennung, wird der Text in Zeichenketten mit bestimmter Länge zerlegt, ohne dass auf die Bedeutung des Ursprungswortes Rücksicht genommen wird (Lewandowski, 2005, 107).

Der **Query Processor** bildet im Aufbau der Suchmaschine die Schnittstelle zwischen Datenbestand und Anwender (Glöggler, 2003, 95). Technisch gesehen ist es Aufgabe des Query Processors, die Suchanfrage des Anwenders zu verarbeiten und eine für den Anwender brauchbare Suchergebnisliste auszugeben. Aufgrund verschiedener Faktoren, wie die Größe des Datenbestandes und der unterschiedlichen Recherchefähigkeiten der Benutzer, entsteht hier eine besondere Herausforderung für Internet-Suchmaschinen als Web Information Retrieval Systeme: Der Query Processor muss aus der Vielzahl der Ergebnisse auswählen und auf Grundlage der Gewichtungsmodelle des Information Retrieval Systems eine Reihung nach Relevanz vornehmen (siehe Kapitel 2.2).

In Anbetracht der Größe des Internets und der Menge an Benutzern, die auf Suchmaschinen zugreifen, müssen Suchmaschinen über eine enorme Rechenleistung und Speichergröße verfügen, damit der Suchdienst in Bezug auf Geschwindigkeit und Dienstgüte alle Benutzer zufrieden stellen kann. Daher arbeiten diese oben genannten Module in mehrfacher Ausführung auf tausenden Rechnern, die um die ganze Welt verteilt sind. Der Marktführer unter den Suchmaschinen, Google, hat laut eigenen Angaben mehr als 10.000 Server in Betrieb, um täglich 150 Millionen Suchanfragen beantworten zu können und 4 Milliarden Dokumente zu indexieren. Allerdings gelten selbst diese imposanten Zahlen als Untertreibung: Laut Garfinkel (2004) hat Google schon im November 2003 mehr als 100.000 Rechner rund um den Globus betrieben und kontrolliert damit das größte Grid-Netzwerk der Welt.

2.2 Suchmaschinen aus der Sicht der Informationswissenschaften

Für die Informationswissenschaften sind Suchmaschinen besonders als Information Retrieval Systeme interessant. Aufgrund des unterschiedlichen Systemumfeldes sind jedoch Suchmaschinen im Internet mit anderen Herausforderungen konfrontiert als klassische Information Retrieval Systeme, zum Beispiel in Fachdatenbanken. Die wichtigsten Herausforderungen an Web Information Retrieval Systeme werden von Lewandowski (2005, 71-76) genannt und lassen sich auf drei Besonderheiten zusammenfassen.

Schon in der Tatsache, dass eine Suchmaschine viele **unterschiedliche Benutzer** bedienen muss, liegen einige Herausforderungen begründet. Schließlich

unterscheiden sich die Nutzer teilweise beträchtlich voneinander, zum Beispiel in Bezug auf das Informationsbedürfnis oder die Recherchekenntnisse. Denn zum einen suchen die Benutzer Informationen zu unterschiedlichen Themen und in unterschiedlichen Detailgraden, andererseits können sie zumeist nicht mit erweiterten Suchfunktionen oder komplexen Abfragesprachen umgehen.

Weiters stellt die **Größe des Internets** eine Herausforderung dar. Einerseits können die Suchmaschinen aufgrund technischer Einschränkungen nicht alle Dokumente, die im Internet verfügbar sind erreichen, andererseits kennt man die genaue Größe des Internet nicht, sondern muss sich bei Angaben bezüglich des Datenbestands auf Schätzungen verlassen. Es entsteht also die Problematik, dass die Suchmaschinen nicht das gesamte Web indexieren und auch keine genaue Angaben zum Grad der Abdeckung machen können.

Letztlich stellen auch die **Besonderheiten der Web-Dokumente** zusätzliche Anforderungen an Suchmaschinen. So müssen die Systeme etwa Dokumente, die in unterschiedlichsten Sprachen verfasst worden sind, verarbeiten und – angefangen bei HTML-Dateien, über PDFs und anderen Textdateien, bis hin zu multimedialen Inhalten, wie Animationen, Bildern sowie Video- und Audio-dateien – mit einer Menge von Formaten umgehen können. Darüber hinaus liegen die Dokumente in unterschiedlichster Qualität, Länge und Granularität vor, sodass die Dokumente bewertet und gegebenenfalls, etwa bei unerwünschten Inhalten (Spam) oder bereits indizierten Ressourcen (Dubletten), aus dem Index entfernt werden müssen. Die schwache Strukturierung insbesondere von HTML-Dateien (Lewandowski, 2005, 59) und die Möglichkeit mit Hyperlinks auf andere Dokumente zu verweisen, erfordern zusätzliche Kriterien für die Inhaltsbewertung.

Aufgrund dieser Herausforderungen, die das Systemumfeld Internet für Suchmaschinen bereit hält, weisen Suchmaschinen zahlreiche Unterschiede zu klassischen Information Retrieval Systemen auf. Dennoch finden bei Web Information Retrieval Systemen die Modelle des klassischen Information Retrievals, wenn auch mit anderen Schwerpunkten, zum großen Teil Anwendung. Einen besonderen Stellenwert hierbei nehmen, besonders in Hinblick auf Suchmaschinen-Marketing, die Verfahren des sogenannten **Relevance Ranking** ein. Denn erst aufgrund der unübersichtlichen Menge an Ergebnissen, die Abfragen in Suchmaschinen meistens zurückliefern, und der überwiegend unzureichenden Recherchekenntnisse der Nutzer, um die gefundenen Ergebnisse einzuschrän-

ken, wird es notwendig, die Ergebnisliste nach Relevanz zu reihen (Lewandowski, 2005, 89).

Laut Aaberg (2006) besteht der Begriff Relevanz in Zusammenhang mit Internet-Suchmaschinen aus einigen Grundprinzipien. Demnach gilt ein Dokument als besonders relevant, wenn es viele Suchbegriffe enthält (statistics), möglichst aktuell ist (freshness), geographischen Bezug zum Suchenden hat (geography) und nicht zuletzt, wenn es qualitativ hochwertig ist (quality) oder von einer anerkannten Kompetenz erstellt worden ist (authority). Um diese schwierige und nicht unumstrittene Relevanzbewertung von Web-Dokumenten durchzuführen, stehen den Suchmaschinen nur eine begrenzte Menge an Kriterien zur Verfügung. Dabei unterscheiden sich die Suchmaschinen kaum mehr darin, welche Kriterien sie zur Bewertung verwenden, sondern vielmehr versuchen sich die Betreiber durch unterschiedliche Gewichtung dieser de-facto Standardfaktoren von der Konkurrenz abzuheben.

Lewandowski (2005, 90) unterscheidet in diesem Zusammenhang alle zur Verfügung stehenden Kriterien in On-The-Page Kriterien und Off-The-Page Kriterien. Mit Hilfe der On-The-Page Faktoren wird bestimmt, wie exakt ein Dokument einer Anfrage entspricht, weshalb man auch von anfrageabhängigen Kriterien spricht. Hier kommen Verfahren des klassischen Information Retrievals zum Einsatz, wenn die Texte statistisch und informationslinguistisch ausgewertet werden, um etwa die Häufigkeit oder die Position eines Begriffes festzustellen. Im Gegensatz dazu versucht man mit den Off-The-Page Faktoren anfrageunabhängige Kriterien zu generieren, um die Qualität und Autorität der Dokumente besser einschätzen zu können. Es wird zum Beispiel versucht, anhand von Nutzungsstatistiken oder der Analyse des Systems von eingehenden und ausgehenden Hyperlinks Rückschlüsse auf die Qualität gewisser Seiten unabhängig von der Anfrage zu ziehen.

Zur Veranschaulichung führt Lewandowski (2005) folgendes Beispiel zur Erklärung der Bedeutung von anfrageabhängigen und anfrageunabhängigen Faktoren für die Suchmaschinen an:

„Würden sie nur anfrageabhängige Verfahren einsetzen, könnten sie nicht zwischen dem Original und einer Kopie bzw. Manipulation eines Dokuments unterscheiden [...]. Der alleinige Einsatz von anfrageunabhängigen Verfahren ist nicht möglich, da dies zur Ausgabe der immer gleichen Trefferliste unabhängig von der Suchanfrage führen würde.“

Eine genaue Betrachtung der Kriterien, auf die man besonders im Rahmen eines Suchmaschinen-Optimierungsprozesses Acht geben muss, folgt in den Kapiteln 3.3 und mit einem Schwerpunkt auf On-The-Page Kriterien im Rahmen der empirischen Untersuchung in Kapitel 4.

2.3 Die Entwicklung des Suchmaschinenmarkts

Die Geschichte von Internet-Suchmaschinen (Michael/Salter, 2003, 1-9) geht zurück bis ins Jahr 1990. Schon zu dieser Zeit, als das Internet hauptsächlich aus FTP-Servern für den Datenaustausch bestand, hilft eine Suchmaschine namens Archie, die im Netzwerk verteilten Dateien zu finden. 1993 werden die Suchmaschinen Veronica und Jughead für die Suche im Gopher-Netzwerk entwickelt. Etwa zur gleichen Zeit durchkämmt der World Wide Web Wanderer das Internet mit dem Ziel, Webserver und URLs automatisiert zu zählen, um Rückschlüsse auf die Größe des World Wide Webs ziehen zu können. Die Entwicklung dieses so genannten Webrobots und die Idee, dass man das gesamte Web indexieren kann, wenn man nur den Verlinkungen folgt, bringt erste Suchmaschinen hervor, die den Titel und die URL von Webseiten speichern. Etwa 1994 geht mit WebCrawler die erste Suchmaschine online, die Webseiten im Volltext indexiert. 1995 werden die Suchmaschinen Lycos, Infoseek und Altavista mit dem Anspruch gegründet, über Indexgrößen in Millionenhöhe zu verfügen. Mit Hotbot startet 1996 eine Suchmaschine, deren Geschwindigkeit es erlaubt, den Index regelmäßig auf den neuesten Stand zu bringen. Inktomi, ein wichtiger Suchmaschinenhersteller, tritt 1996 in den Markt ein und versorgt eine Reihe von großen Portalen mit seiner Suchtechnologie. Google, 1998 gegründet, setzt sich am Markt gegen Inktomi ebenso wie gegen die schon 1997 gegründeten Suchmaschinenhersteller FAST (Fast Search & Transfer) und Teoma durch.

Seither versuchen zahlreiche Suchmaschinenanbieter, wie MSN Search, Yahoo!, Ask Jeeves oder Wisenut, die Marktführerschaft von Google zu brechen – jedoch bisher erfolglos, denn die Entwicklung geht eher in die Richtung, dass Google seine Marktanteile weiter ausbaut (siehe Abbildung 2).

Aus den Diagrammen auf Seite 10 erkennt man auch, dass der Anteil anderer Suchmaschinen immer kleiner wird. Der Markt konzentriert sich also auf drei Anbieter, nämlich Google, Yahoo! und MSN. Viel stärker fällt diese Kon-

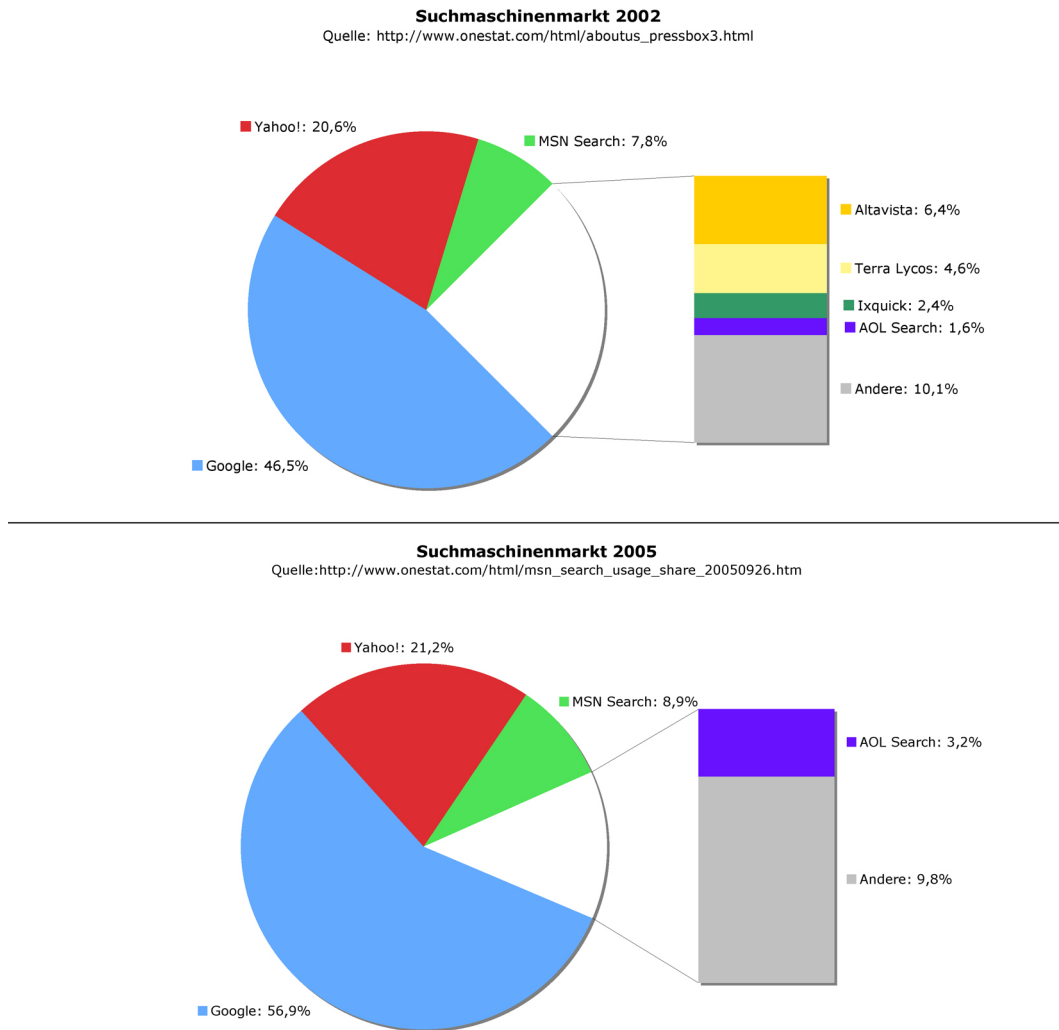
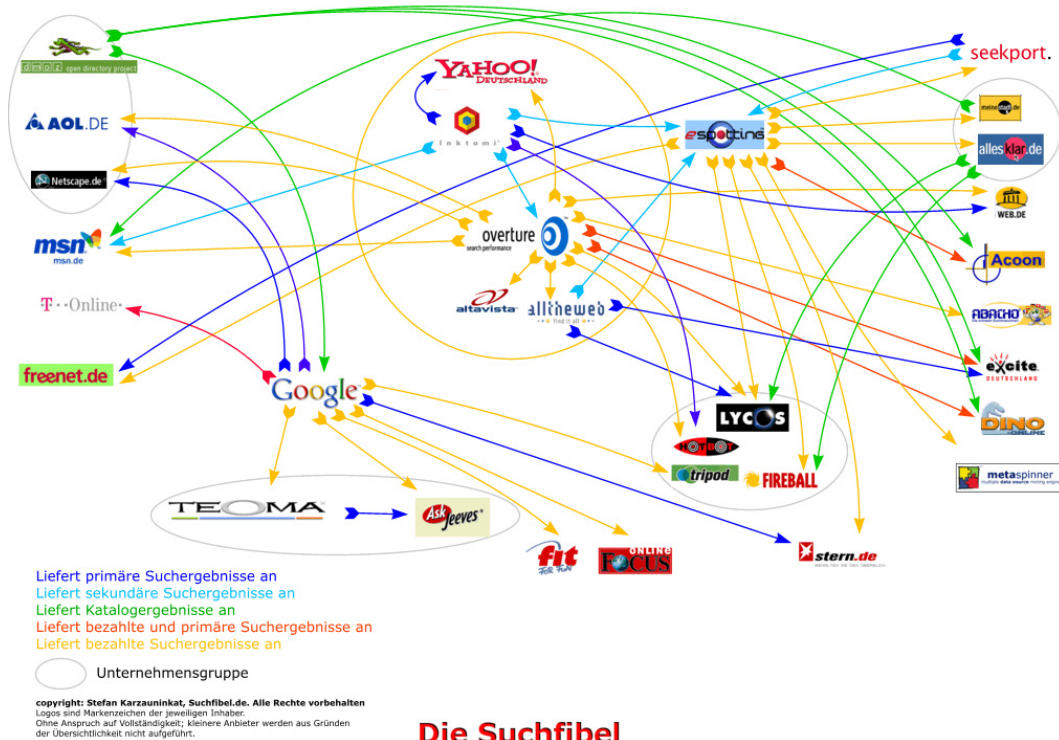


Abbildung 2: Anteile von Google, Yahoo! und MSN Search am Suchmaschinenmarkt in den Jahren 2002 und 2005 (Quelle: OneStat).

zentration auf, wenn man sich die Verflechtungen der Suchmaschinenbetreiber in Deutschland (siehe Abbildung 3) und international (siehe Abbildung 4) ansieht. Außer Google, das ein erstaunlich großes Netzwerk von Partnerschaften betreibt, fällt hier vor allem die Unternehmensgruppe um Yahoo! ins Auge: Neben dem ursprünglichen Verzeichnisdienst hat Yahoo! den Suchmaschinentechnologieanbieter Inktomi, die beiden bekannten Suchmaschinen Altavista und AllTheWeb und die Paid Placement Suchmaschine Overture⁵ akquirieren können und versorgt weltweit eine Menge an Portalen mit klassischen und bezahlten Suchergebnissen. Das dritte Unternehmen, von dem viele

⁵ Gegründet unter dem Namen GoTo, gilt Overture als erste Suchmaschine, die nur bezahlte Ergebnisse in ihren Resultaten angezeigt hat. Mittlerweile, als Folge der Übernahme durch Yahoo!, heißt das Unternehmen Yahoo! Search Marketing.

Beziehungsgeflecht der Suchdienste in Deutschland und international, Stand 17.12.2004



Die Suchfibel

Abbildung 3: Die Vernetzung der Suchmaschinen in Deutschland, zusammengestellt von Karzauninkat (2002).

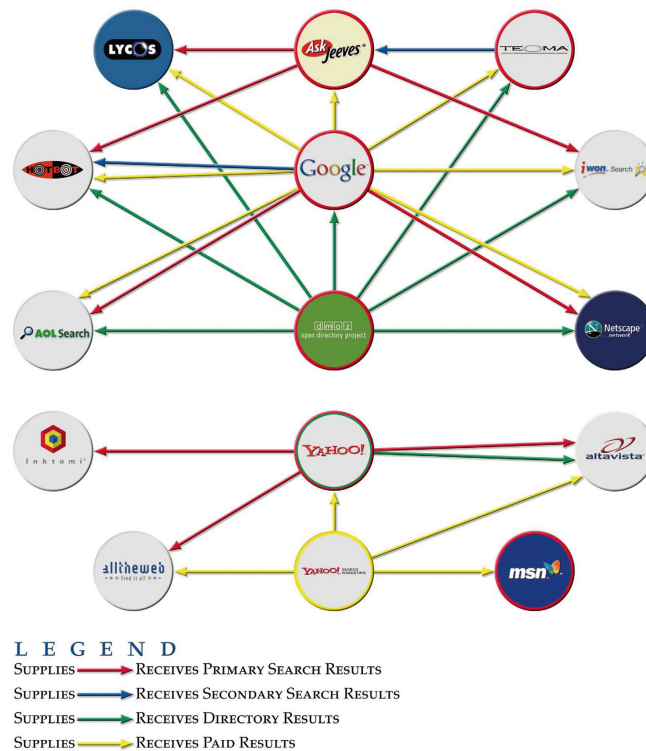


Abbildung 4: Die Vernetzung der Suchmaschinen weltweit, zusammengestellt von der Bruce Clay Inc. (2006).

Verbindungen in der Grafik von Karzauninkat (2002) ausgehen, hieß bis vor Kurzem noch Espotting. In der Zwischenzeit wurde es von der FindWhat-Unternehmensgruppe übernommen und in MIVA umbenannt.

Sowohl in Abbildung 3 von Karzauninkat (2002), als auch in Abbildung 4 von Bruce Clay Inc. (2006) fällt auf, dass MSN, das Portal von Microsoft, von Yahoo!, also einem der großen Konkurrenten im Internet, die bezahlten Ergebnisse bekommt. Das Ende dieser Zusammenarbeit ist jedoch schon absehbar, denn seit Monaten entwickelt Microsoft am sogenannten adCenter, einem eigenen System, um Werbung auf den Suchergebnisseiten zu integrieren, und seit September 2005 wird es auf den Suchseiten von Frankreich und Singapur getestet.

3 Suchmaschinen-Marketing

Für die Betreiber der Suchmaschinen ist Suchmaschinen-Marketing mittlerweile mehr, als nur eine Möglichkeit, den Betrieb und die Weiterentwicklung eines Suchdienstes zu finanzieren. Einer Schätzung von Piper Jaffay (zitiert in Battelle 2005, 34) haben die Suchmaschinen mit bezahlten Ergebnissen etwa 4 Milliarden Dollar eingenommen und die Einnahmen sollen bis 2010 23 Milliarden Dollar erreichen. Tatsächlich macht zum Beispiel Google schon heute 95 Prozent seines Umsatzes mit Werbung⁶ (Battelle, 2005, 164) und nur etwa ein Prozent mit Lizenzen, Suchmaschinentechnologie und Services (Speck, 2006).

Vor Jahren wehrte sich die Suchmaschine Altavista noch gegen eine kommerzielle Vermarktung der Suchergebnisse, indem sie in ihren Richtlinien feststellten, „*This is a search engine, not a promotion engine!*“ (zitiert in Karzauninkat 2002, 189). Zu diesem Zeitpunkt richtete sich der Satz vor allem auch gegen jene Webmaster, die Methoden herausgefunden haben, um den Suchmaschinenalgorithmus zu manipulieren und die Platzierungen beeinflussten. Diese Seitenbetreiber hatten angesichts der steigenden Bedeutung von Suchmaschinen als Einstiegspunkt für Internetsuchen folgendes erkannt: „*Schafft ein Webmaster es, dass Suchmaschinen die eigene Site an prominenter Stelle*

⁶ Diese Angabe umfasst sowohl die bezahlten, keywordbezogenen Ergebnisse auf eigenen oder fremden Suchergebnisseiten (Google AdWords), als auch webseitenbezogene Werbung (Google AdSense), die in dieser Arbeit nicht besprochen werden.

der Ergebnislisten präsentieren, so ist seiner Online-Präsenz ein breiter Besuchersstrom sicher [...]“ (Bager, 2005a).

Das Bestreben der Webmaster, Seiten so zu erstellen, dass Suchmaschinen sie als möglichst hochwertig erachten und hoch in den Ergebnislisten ausgeben, und die Idee der Suchmaschinenbetreiber, bestimmte Bereiche der Suchmaschinenergebnislisten zu reservieren und an den Höchstbietenden zu verkaufen, bilden die beiden Grundsteine für Suchmaschinen-Marketing, so wie wir es heute kennen.

3.1 Definition und Ziel von Suchmaschinen-Marketing

Der Definition des Arbeitskreises Suchmaschinen-Marketing im Bundesverband Digitale Wirtschaft (2004) folgend, umfasst Suchmaschinen-Marketing *„alle Maßnahmen zur Gewinnung von qualifizierten Besuchern über Suchergebnisseiten in Suchmaschinen.“* Ziel des Suchmaschinen-Marketing muss es daher sein, eine Webseite langfristig bei möglichst allen relevanten Suchphrasen auf der ersten Ergebnisseite der Suchmaschinen zu platzieren und so potentielle Interessenten von der Suchmaschinenergebnisseite auf die eigene Webseite zu führen. Wie Abbildung 5 zeigt, bietet der Aufbau der Suchergebnisseiten zwei Möglichkeiten um eine Webseite erfolgreich zu platzieren. Einerseits kann man eine Webseite unter den klassischen Ergebnissen, andererseits unter den bezahlten Einträgen platzieren.

Um die Platzierung einer Webseite unter den klassischen Ergebnissen zu erhöhen, bedient man sich des Instruments der **Optimierung**⁷. Um diese Verbesserung der Position herbeizuführen, kann man sich verschiedener Methoden bedienen, damit die Suchmaschinen eine Webseite bei einigen wenigen Suchphrasen höher unter den klassischen Ergebnissen der Suchergebnislisten reiht. Zum einen versucht man Inhalt und Aufbau einer Webseite zu verbessern, um von der Suchmaschine für gewisse Kombinationen von Suchbegriffen als relevant erachtet zu werden (siehe Kapitel 3.3.1). Zum anderen muss man auch abseits der Seite dafür zu sorgen, dass die Suchmaschine von der Qualität und Autorität der Inhalte überzeugt wird (siehe Kapitel 3.3.2). Grundsätzlich fallen von Seiten des Suchmaschinenbetreibers keine Kosten an, wenn ein Webmaster einen Internetauftritt optimiert. Nichtsdestotrotz kann die Optimierung

⁷ Wird häufig auch Suchmaschinenoptimierung, Websiteoptimierung, SEO,... genannt.

The screenshot shows a Yahoo! search results page for the keyword "blumen". The page layout includes a search bar at the top, navigation links (Web, Bilder, Video, Nachrichten), and search filters (Suche, Das Web, Seiten auf Deutsch, Seiten aus Österreich). The search results are displayed in a list format, with the first result being "Blume 2000 - Blumen Versand Online - Geschenkideen u. Blumengrüße". To the right of the search results, there is a section titled "SPONSOREN-LINKS" which contains several sponsored links, including "Blumen top günstig bei Quelle.at bestellen" and "Blumen mit Fleurop versenden, online bestellen". A red box highlights the sponsored links section, and a green box highlights the main search results section. The text "bezahlte Einträge" (paid entries) is written in red over the sponsored links, and "klassische Ergebnisse" (classic results) is written in green over the main search results.

Abbildung 5: Klassische Ergebnisse und bezahlte Einträge am Beispiel der Ergebnissseite der Suchmaschine Yahoo!

eines Internetauftritts einen durchaus großen zeitlichen Aufwand bedeuten, zum Beispiel für die Aneignung der notwendigen fachlichen Kenntnisse oder für die Planung, die Durchführung und die Administration einer Kampagne. Schließlich bedeutet Optimierung nichts anderes, als langfristig von den Suchmaschinen problemlos gefunden und erfolgreich indexiert zu werden, um zu bestimmten Suchbegriffen möglichst hoch gereiht wiedergefunden zu werden.

Das zweite Instrument des Suchmaschinen-Marketing ist es, **keywordbezogene Werbung**⁸ unter den bezahlten Einträgen zu schalten. Hierbei erkauft man sich zu einer beliebigen Anzahl an Suchbegriffen, sogenannten Keywords, eine Platzierung in jenem Bereich, den der Suchmaschinenbetreiber für bezahlte Einschaltungen reserviert hat. Die beste Position wird in einem Auktionsmodus an jene Person vergeben, die bereit ist am meisten für einen Klick auf seine Anzeige zu bezahlen. In diesem so genannten pay-per-click-Verfahren (PPC)

⁸ Besonders im Internet finden sich zahlreiche andere Bezeichnungen für Werbung im Sinne von Suchmaschinen-Marketing, zum Beispiel keyword advertising, pay-per-click advertising (PPC), cost-per-click advertising (CPC), paid placement, pay-for-placement und viele andere mehr. Eine einheitliche Übersetzung ins Deutsche gibt es nicht.

entstehen nämlich erst dann Kosten für den Werbetreibenden, wenn jemand explizit von der Einschaltung Notiz nimmt und darauf klickt⁹. In der Regel bezahlt man für einen Klick, abhängig von der Konkurrenz, wenige Cent, bei gewissen Begriffen und Branchen (denkbar etwa bei Immobilien, Glücksspiel, ...) allerdings können die Preise durchaus deutlich über einen Euro hinaus gehen. Den Kosten, die durch Suchmaschinen-Werbung entstehen können sind also, sofern man kein maximales Tagesbudget für die Kampagne angibt, keine Grenzen gesetzt. Nichtsdestotrotz hängt jedoch der langfristige Erfolg von Werbung auf den Ergebnislisten der Suchmaschinen weniger von Geld, als von einer guten Planung der Kampagne ab.

Suchmaschinen-Marketing ist also der Überbegriff für Optimierung und Werbung¹⁰ und zählt innerhalb des Marketing, wie alle Online Marketing Instrumente, zum Direktmarketing. Diese Instrumente sind *„nicht-persönliche Kommunikations- und Kontaktmittel, durch die gezielt mit ausgesuchten Kunden und potentiellen Kaufinteressenten kommuniziert wird, um bei ihnen eine Reaktion auszulösen.“* (Kotler/Bliemel, 2001, 882 bzw. 1185-1230). Bei Suchmaschinen-Marketing tritt man, völlig gegensätzlich zur klassischen Werbung, wie zum Beispiel Anzeigen in Zeitungen, Zeitschriften, Fernsehen, Radio und Kino, über die Suchmaschinenergebnislisten direkt mit Menschen in Kontakt, die potentiell hohes Interesse an der beworbenen Website haben – sogenannte Streuverluste fallen daher gering aus. Dabei besteht bei Suchmaschinen-Marketing-Kampagnen die besondere Herausforderung, die Keywords so zu wählen, dass die beworbene Website gut von allen potentiellen Interessenten – und möglichst nur von potentiellen Interessenten – gefunden werden kann, und somit für den Interessenten das Gefühl entsteht, dass er durch seine Suche (insbesondere mit der Wahl der Suchbegriffe) mit dem Werbetreibenden in Kontakt getreten ist, und nicht der Werbetreibende mit ihm.

Der Definition von Direktmarketing folgend, muss auch eine Suchmaschinen-Marketing-Kampagne, als ein Teil im Marketing-Mix einer Organisation, immer das Ziel haben, vom Interessenten eine Reaktion zu bekommen. Diese Reaktion muss zwar nicht immer der Online Verkauf sein, jedoch zumindest beinhalten, den Interessenten aus der Anonymität des Internet zu holen, also persönliche

⁹ Im Unterschied dazu bezahlt man in der Bannerwerbung und auch in der klassischen Werbung, etwa bei einem Zeitungsinserat, für sogenannte *impressions*, also dafür, wie vielen Personen die Werbung angezeigt wird.

¹⁰ Optimization and advertising are two methods within search engine marketing. (Seda, 2004, 4)

Daten und Angaben vom Benutzer zu erhalten. Schließlich kann nur in den seltensten Fällen *traffic*, also viele Besucher auf der Website zu bekommen, das alleinige Ziel von Suchmaschinen-Marketing sein.

3.2 Suchmaschinen-Werbung

Die scheinbar einfachste Methode, Suchmaschinen-Marketing zu betreiben, ist das Schalten von Anzeigen in jenen Bereichen der Suchergebnislisten, die für bezahlte Ergebnisse reserviert sind. Es steckt jedoch mehr hinter erfolgreicher Suchmaschinen-Werbung, als auf den ersten Blick erkennbar ist. Der Erfolg einer Kampagne im Bereich Suchmaschinen-Werbung hängt vor allem von der Planung der Kampagne im Vorfeld ab. Noch bevor man sich mit der Auswahl der Suchmaschinen und der Keywords beschäftigt und Anzeigentexte und Landing Pages gestaltet, sollte man sicherstellen, dass Suchmaschinen-Werbung überhaupt in den Marketing Mix des Unternehmens passt. Befolgt man die Anweisungen, die Seda (2004, 21 ff.) gibt, sollte man sich die Ziele des Unternehmens (mission statement), die Zielgruppe (target market), den Wettbewerbsvorteil (unique selling point) und die Ziele der Kampagne (primary goals, bzw. call to action) vor Augen führen und niederschreiben. Während die Unternehmensziele, die Zielgruppe und der Wettbewerbsvorteil dabei helfen, die Kampagne richtig auszurichten (Keywordauswahl, Zielgruppenansprache, geographische Einschränkungen, ...), bilden die finanziellen und pragmatischen Ziele der Kampagne die Grundlage für die spätere Erfolgskontrolle (Details siehe Kapitel 3.4).

3.2.1 Pay-Per-Click Suchmaschinen

Auf Basis der Ergebnisse der oben umrissenen Planung werden jene Anbieter von Suchmaschinen-Werbung ausgewählt, mit denen man am besten die Zielsetzungen der Kampagne, insbesondere in Bezug auf Zielgruppe und geographische Einschränkungen, erreichen kann. Dabei wird man nur in den seltensten Fällen auf die großen Anbieter von Suchmaschinen-Werbung, wie Google AdWords, Yahoo! Search Marketing oder Miva, verzichten können. Diese Pay-Per-Click-Suchmaschinen (PPC) haben die größten und bedeutendsten Suchmaschinen und Portale als Partner in ihrem Netzwerk und integrieren die bezahlten Einschaltungen in deren Suchergebnislisten (siehe Kapitel 2.3 und

Abbildungen 3 und 4). Mit dieser weiten Verbreitung der Anzeigen steigt auch die Wahrscheinlichkeit, dass die beworbene Webseite oft besucht wird. Wie schon oben erwähnt, ist jedoch kaum die Zahl der Besucher (traffic) ausschlaggebend für den Erfolg einer Suchmaschinenkampagne, sondern vielmehr, ob sie mit der Organisation in Kontakt treten (conversion).

Anbieter	Regionale Abdeckung
Google AdWords	weltweit
Yahoo! Search Marketing	weltweit
Miva	weltweit
MSN AdCenter	dzt. Frankreich und Singapur
Mirago	Westeuropa mit Schwerpunkt Großbritannien und Deutschland
Eniro	Skandinavien und Baltikum, tw. Deutschland und Polen
QualiGo	Österreich, Deutschland, Schweiz und Niederlande
eTarget	Tschechische Republik, Slowakei Ungarn, Slowenien
GBG	Bulgarien
Onet	Polen
Yandex	GUS, insbesondere Russland
Begun	GUS, insbesondere Russland

Tabelle 1: Auswahl von PPC-Anbietern, die in Europa tätig sind. Regionale Abdeckung laut Anbieterwebseiten.

Zusätzlich zu den großen Anbietern lohnt es sich daher oft auch, in kleineren, regional beschränkten PPC-Suchmaschinen Anzeigen zu schalten. Denn auch wenn die Systeme der großen Anbieter eine lokale Einschränkung auf Länder oder Sprachen zulassen, erreichen sie nicht kleine, geographisch oder fachlich begrenzte Suchmaschinen. Um auch diese Suchmaschinen abzudecken, findet man im Internet kleinere PPC-Suchmaschinen, die Suchmaschinen von besonderer regionaler Bedeutung mit bezahlten Einschaltungen versorgen.

3.2.2 Auswahl der Keywords

Ebenso zentral wie die Auswahl der PPC-Suchmaschinen, ist die Auswahl der Keywords für den Erfolg von Kampagnen im Bereich Suchmaschinen-Werbung. Schließlich hängt es von den Keywords ab, ob der bezahlte Eintrag angezeigt

wird, oder nicht. Die Herausforderung hierbei besteht darin, nicht zu wenige, aber auch nicht zu viele Begriffe und Begriffskombinationen auszuwählen. Schließlich wird man einerseits von vielen potentiellen Kunden nicht gefunden, wenn die Anzeigen nicht zu allen relevanten Begriffen erscheinen, andererseits wird Werbebudget durch unzutreffende Keywords verschwendet, weil sie zwar die Besucherzahlen auf der Website erhöhen, also traffic generieren, jedoch die Ziele der Kampagne nicht erfüllen. Eine regelmäßige Kontrolle der Kampagne (siehe Kapitel 3.4) liefert dabei wichtige Informationen über den Erfolg der gewählten Begriffe. Innerhalb kurzer Zeit kann man also falsch gewählte Begriffe identifizieren und eliminieren und erfolgreiche Keywords weiter ausbauen.

Einen Hinweis, wie die Keywords und Keywordkombinationen zu wählen sind, geben einige Statistiken: Seda (2004, 43) erwähnt eine Statistik von OneStat, wonach mit 78,31 Prozent der überwiegende Teil der Suchmaschinenbenutzer Suchphrasen mit 3 oder weniger Wörtern bildet¹¹. Battelle (2005, 27f.) stellt zusätzlich dazu fest, dass nur wenige Suchbegriffe sehr oft gesucht werden. Im Gegensatz dazu werden täglich zwischen 50 und 100 Millionen¹² einzigartige Suchanfragen abgeschickt, was einen Anteil von bis zu 50 Prozent bedeutet. Daraus kann man schließen, dass sich Begriffskombinationen mit bis zu drei Wörtern lohnen und dass bei der Auswahl durchaus Kreativität gefragt ist, um bei möglichst vielen relevanten Suchanfragen, auch wenn sie nur ein Mal pro Tag gesucht werden, gefunden zu werden.

Für die Erstellung der Keywordliste sollte man sich zuallererst die sogenannten **Branded Keywords** sichern (Seda, 2004, 28). Branded Keywords sind alle markenrechtlich geschützten Bezeichnungen und Beschreibungen eines Unternehmens und deren Produkte. Diese Keywords haben einerseits den Vorteil, dass sie um das jeweilige Mindestgebot der PPC-Suchmaschine zu kaufen sein müssen, da sie ausschließlich nur vom Inhaber der Markenrechte registriert werden dürfen. Andererseits gelten diese Begriffe als sehr erfolgreich, weil viele Menschen ein Produkt, eine Dienstleistung oder ein Unternehmen nach dessen Markennamen in der Suchmaschine suchen. Dazu zählen weiters Domainnamen und URLs. Neben Seda (2004, 30) spricht auch Krug (2002) vom Phä-

¹¹ Die Statistik ist unter http://www.onestat.com/html/aboutus_pressbox20.html im Internet abrufbar. Mit 29,22 Prozent wird am häufigsten mit zwei Wörtern gesucht. Lewandowski (2005, 37) spricht, basierend auf einer anderen Untersuchung, von durchschnittlich 2,6 Termen je Anfrage.

¹² Die Spanne ergibt sich aus dem Unterschied zwischen einer Studie von Piper Jaffray (50 Mio.) und den Angaben von Google (100 Mio.).

nomen, dass eine unvorstellbar große Menge von Internetbenutzern URLs von Websites nicht direkt in die Adresszeile des Browsers eingeben, sondern in die Eingabemaske einer Suchmaschine und schließlich die Webseite über die Suchergebnisse ansurfen.

Der größere Teil der Keywordliste wird in den meisten Fällen aus **Generic Keywords** bestehen. Hier versucht man zu antizipieren, welche Begriffe ein Suchmaschinennutzer benutzt, wenn er das Angebot, das auf der beworbenen Internetseite bereit gestellt wird, sucht – auch wenn er noch nicht weiß, dass es dieses Angebot überhaupt gibt. Zu diesem Zweck bedient man sich jener Wörter, die nicht nur organisationsintern für die Beschreibung des Unternehmens und dessen Tätigkeit benutzt werden, sondern vor allem auch jene Bezeichnungen, die Kunden verwenden, wenn sie etwa telefonisch oder schriftlich mit der Organisation in Kontakt treten. Auch brancheninterne Begriffe und anderer Fachjargon bieten sich hier an.

Um die Keywordliste in einem ersten Schritt zu erweitern, empfiehlt es sich, Variationen der bereits gefundenen Branded Keywords und Generic Keywords als Suchbegriffe auszuprobieren. Hierzu zählen unter anderem Kombinationen aus mehreren Begriffen, Abwandlungen der Begriffe zu Verben oder Adjektiven, Pluralformen, Rechtschreibfehler und Auslassungen. Darüber hinaus bieten die PPC-Suchmaschinen, aber auch Drittanbieter, Werkzeuge an, die anhand von bisher gesuchten Phrasen Suchbegriffe vorschlagen¹³. Als zusätzliche Quelle bei der Suche nach passenden Begriffen und Begriffskombinationen gelten außerdem die Suchmaschinen-Marketing Aktivitäten der Mitbewerber und User Befragungen.

Welche Begriffe man wählt und woher man sie bekommt, bleibt zuletzt immer der Kreativität der Werbetreibenden überlassen. Man muss jedoch immer im Hinterkopf behalten, dass eine möglichst lange Keywordliste nicht unbedingt bedeutet, dass die Kampagne besonders erfolgreich ist. Vielmehr wird es die Besucher frustrieren, wenn sie auf den beworbenen Seiten nicht genau das finden, wonach sie – nicht zuletzt durch die Wahl ihrer Suchbegriffe – gesucht haben.

¹³ Marktführer Google bietet zum Beispiel das sogenannte Google AdWords Keyword Suggestion und Yahoo! Search Marketing das Overture's Search Term Suggestion Tool.

3.2.3 Formulierung von Anzeigentext und Landing Page

Ein Element, das also zum Erfolg von Suchmaschinen-Werbung beiträgt, ist die richtige Auswahl der Keywords, die bewirkt, dass die Anzeige bei allen interessanten Suchanfragen aufscheint. Um jedoch potentielle Interessenten von den Suchmaschinenergebnissen auf die beworbene Webseite zu führen, ist es notwendig, dass der Suchende auf die Textanzeigen klickt. Denn solange niemand auf die Einschaltung klickt, entstehen nicht nur keine Kosten, sondern auch noch kein Nutzen. Die Gestaltung des bezahlten Eintrages stellt daher ein weiteres kritisches Element bei Suchmaschinen-Werbung dar.

Unabhängig von der Pay-Per-Click Suchmaschine folgt der Aufbau der bezahlten Anzeigen, in Anlehnung an die klassischen Ergebnisse, immer dem gleichen Muster. Die Anzeigen bestehen immer aus (1) einem Titel, der sich grafisch von der restlichen Anzeige unterscheidet, (2) einer kurzen Beschreibung und (3) einer URL zur sogenannten Landing Page, also einer ausgewählten Seite der beworbenen Website. Für jedes der drei Teile der Anzeige stehen aber nur eine begrenzte Anzahl von Zeichen zur Verfügung.

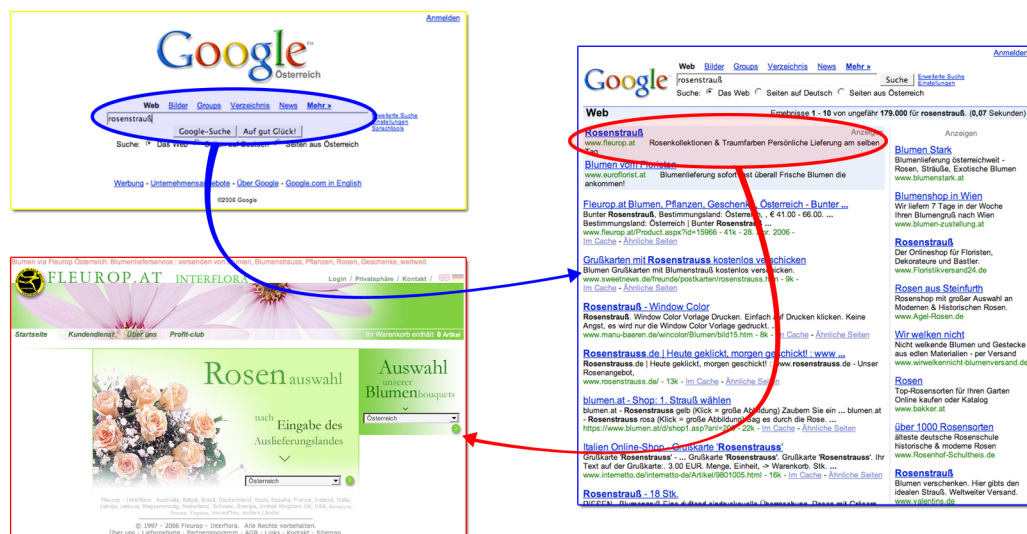


Abbildung 6: Die Beziehung von Keywords, Textanzeige und Landing Page.

Die Herausforderung für die Gestalter der Anzeigentexte besteht nun darin, diesen knapp bemessenen Platz am Besten auszunützen ohne jedoch gegen die Richtlinien der Suchmaschinenbetreiber zu verstoßen. Es wird nämlich jede Anzeige einer redaktionellen Prüfung durch den Suchmaschinenbetreiber unterzogen, bevor sie geschaltet wird. Eigens dafür angestellte Mitarbeiter sor-

gen dafür, dass der Suchmaschinenbenutzer nicht belästigt wird. So sind in den Anzeigentexten keine Blockbuchstaben oder Superlative erlaubt, es dürfen keine Telefonnummern oder URLs in der Beschreibung vorkommen und mit Rufzeichen, aber auch dem kaufmännischen Und, muss sparsam umgegangen werden. Darüber hinaus müssen Keyword und Anzeige zum Angebot auf der Landing Page passen, ebenso, wie sichergestellt sein muss, dass die Zielseite keine Pop-Ups enthält, oder den User daran hindert einfach (z. B. über den Zurück-Button des Browsers) wieder auf die Suchergebnisliste zurückzukehren.

Um gegenüber den anderen bezahlten Einschaltungen einen Vorteil zu haben, kann man sich also keiner Tricks bedienen. Vielmehr gilt es, den Suchenden durch den Anzeigentext zu überzeugen. Wichtig, laut Seda (2004, 48ff.), ist es, die Zielgruppe im Auge zu behalten, indem man zum Beispiel die Suchworte des Suchmaschinennutzers im Anzeigentext verwendet und ihn zu einer Handlung auffordert, oder versucht, den Vorteil des Angebots herauszustreichen, Problemlösungen anzubieten oder den User etwa durch ein zeitlich begrenztes Angebot oder ein Incentive auf die Seite zu locken. Wie auch immer der Anzeigentext gestaltet ist, Rechtschreibfehler und tote Links zur Landing Page sollte man selbstverständlich auf alle Fälle vermeiden!

Hat man es schließlich geschafft, dass der Suchmaschinennutzer auf die Anzeige klickt, führt ihn der angegebene Link auf die beworbene Website. Auf dieser Website sind für das Suchmaschinen-Marketing insbesondere die sogenannten Landing Pages interessant. Denn diese Seiten sind der Einstiegspunkt für alle Besucher, die von den bezahlten Einschaltungen in den Suchergebnislisten kommen. Von den Landing Pages hängt es nun ab, ob man den potentiellen Besucher, der schon mit einem relevanten Keyword gesucht hat und sein Interesse mit einem Klick auf die Anzeige bekundet hat, als Kunde gewinnen kann. Hier entscheidet sich für jeden Besucher, ob die Investition – durch den Klick auf die Anzeige sind ja Kosten entstanden – sinnvoll war.

Um daher nicht nur Besucher, sondern vor allem Kunden¹⁴ auf die Website zu führen, gilt es, den Landing Pages eine besondere Beachtung zu schenken. Es ist dabei nicht wichtig, ob und wie die Landing Pages in die Website integriert

¹⁴ Seda (2004, 61) nennt das Kapitel, das sich mit der Bedeutung der Landing Page für die Kampagne beschäftigt „*Landing Pages That Convert Visitors into Buyers*“.

sind¹⁵, sondern vielmehr, dass die Suche des Benutzers genau auf der beworbenen Seite zu einem erfolgreichen Ende kommt, weil der Besucher sofort genau das findet, was er gesucht hat.

Um dieser Anforderung zu genügen, muss sich der Besucher vom Text der Landing Page angesprochen fühlen. Der Text sollte auf die Anzeige Bezug nehmen und sowohl den Verstand, als auch die Gefühle des Besuchers ansprechen, ohne unseriös zu wirken. In Bezug auf das Layout muss sichergestellt sein, dass der Besucher das Ziel seines Besuches schnell und einfach erreicht. Zu diesem Zweck müssen die wichtigsten Informationen, Links und Buttons auf den ersten Blick erkennbar sein, sodass sich die Aufmerksamkeit des Besuchers fokussiert. Dabei spricht man häufig davon, dass diese besonders wichtigen Informationen *above-the-fold*, also ohne scrollen zu müssen, sichtbar sein müssen und dass man sogar Elemente (z. B. Teile der Navigation), die dem Besucher bei der Erfüllung seiner Aufgabe nicht unmittelbar helfen können, ausblenden sollte, um diese Fokussierung noch zu verstärken.

3.2.4 Gefahren und Risiken beim Einsatz von Paid Placements

Eine sorgfältige Planung der Kampagne erspart dem Werbetreibenden eine Reihe von Problemen, die bei der Durchführung von Suchmaschinen-Werbung auftreten können. Gewisse Risiken bestehen jedoch trotz noch so guter Planung. So zum Beispiel die Gefahr des Klickbetrugs: Zusammengefasst bedeutet Klickbetrug, dass ein Konkurrent – absichtlich und selbstverständlich ohne Kaufabsicht – bezahlte Einschaltungen anklickt, in der Absicht, der Organisation, die Suchmaschinen-Werbung betreibt, finanziellen Schaden zuzufügen. Natürlich versuchen die Suchmaschinenbetreiber betrügerische Aktivitäten zu erkennen, man sollte jedoch auch selbst die Systeme, die einem zur Erfolgskontrolle zur Verfügung stehen, nach verdächtigen Mustern durchsehen und den Suchmaschinenbetreiber auf Unregelmäßigkeiten und Auffälligkeiten aufmerksam machen.

Ebenso wichtig ist es, nach markenrechtlichen Verstößen Ausschau zu halten. Obwohl die Suchmaschinenbetreiber die Anzeigen prüfen, kommt es vor, dass jemand eine rechtlich geschützte Marke als Keyword anmeldet oder in einer

¹⁵ Viele Landing Pages sind von der normalen Website aus nicht erreichbar und vor der Indizierung durch Suchmaschinen mit Hilfe einer robots.txt-Datei (siehe Kapitel 3.3.3) geschützt.

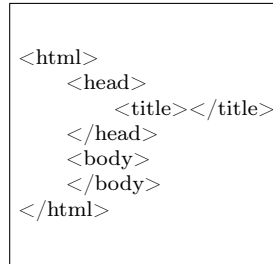
Anzeige oder auf einer Landing Page benutzt. Hier muss man wiederum selbst immer wieder die Ergebnislisten überprüfen und den Betreiber des Suchdienstes gegebenenfalls auf den rechtlichen Verstoß hinweisen.

Letztlich gilt auch für Suchmaschinen-Werbung das Problem, mit denen alle Direktmarketinginstrumente zu kämpfen haben, nämlich dass sich Kunden „*belästigt, übervorteilt, schlimmstenfalls betrogen sowie in ihrer Privatsphäre verletzt*“ fühlen (Kotler/Bliemel, 2001, 1226). Daher sinkt die Akzeptanz von Direktmarketingkampagnen, unter anderem auch von Suchmaschinen-Anzeigen.

3.3 Suchmaschinen-Optimierung

Im Wesentlichen bezeichnet der Begriff Suchmaschinen-Optimierung die Gesamtheit aller Maßnahmen, um eine Website möglichst hoch unter den klassischen Ergebnissen der Ergebnislisten zu platzieren. Analog zu technischen Komponenten einer Suchmaschine und deren Aufgabe (siehe Kapitel 2.1) haben diese Maßnahmen das Ziel, dass die Webseite fehlerfrei vom Webroboter-System der Suchmaschinen gefunden wird, vom Information Retrieval System als qualitativ hochwertig eingeschätzt wird und vom Query Processor der Suchmaschine zu möglichst allen relevanten Suchbegriffen unter den ersten Suchergebnissen ausgegeben wird. Um hier langfristig Erfolg zu haben, muss die Seite möglichst vielen Faktoren, die die Suchmaschine als Qualitätsmerkmal für die Relevanzbewertung heranzieht, entsprechen. Aufgrund der Vielzahl der Faktoren – Google (2005b) behauptet, dass die Reihenfolge von über 100 Faktoren bestimmt wird – und der spärlichen Informationen seitens der Suchmaschine über deren Gewichtung, muss man sich bei der Auswahl der Methoden auf einschlägige Literatur und persönliche Erfahrungen verlassen.

Wie schon in Kapitel 2.2 erwähnt, unterscheidet man die Faktoren, die die Suchmaschinen für die Erschließung der Web-Dokumente verwendet, grundsätzlich in die beiden Gruppen On-The-Page und Off-The-Page Faktoren. Analog dazu bezeichnet Glöggler (2003, 115ff. und 169ff.) Maßnahmen, die zu einer Verbesserung der Relevanzbewertung durch die Suchmaschine führen, als On-The-Page, beziehungsweise Off-The-Page Methoden. Die von einschlägigen Autoren oftmals genannten Maßnahmen sind in den kommenden beiden Abschnitten erwähnt. Darüber hinaus wird sich Kapitel 4 mit der Wirksamkeit der genannten Methoden je nach Suchmaschinenanbieter beschäftigen.



```
<html>
  <head>
    <title></title>
  </head>
  <body>
  </body>
</html>
```

Abbildung 7: Grundsätzlicher Aufbau einer HTML-Datei.

3.3.1 On-The-Page Methoden

Als On-The-Page Methoden gelten alle Verfahren, die direkt auf das Dokument angewendet werden. Dabei liegt das Augenmerk auf der Wortwahl, Platzierung und Gestaltung der Texte und Inhalte. Thurow (2003, 18) spricht in diesem Zusammenhang von der Textkomponente (text component) der Suchmaschinen-Optimierung. Der Webmaster hat also in Hinblick auf die Optimierung freie Hand bei der Anwendung der Maßnahmen und ist nicht auf die Hilfe anderer Webmaster, der Benutzer oder gar der Suchmaschine angewiesen. Primäres Ziel dieser Verfahren ist es, die Web-Dokumente so zu erstellen, dass sie zu relevanten Suchbegriffen gefunden werden.

Textgestaltung und Keyworddichte sind jene Faktoren eines Dokuments, das von den Suchmaschinen mit klassischen informationswissenschaftlichen Verfahren ausgewertet werden. Vergleichbar mit klassischen Information Retrieval Systemen werden zum Zwecke der automatischen Erschließung Kriterien wie Worthäufigkeiten und Position der Suchbegriffe im Dokument herangezogen (Lewandowski, 2005, 90ff.). Die Suchmaschine leitet die Deskriptoren für eine Webseite davon ab, wo und wie oft die Wörter vorkommen.

Aus Abbildung 7, die den grundsätzlichen Aufbau einer HTML-Seite (Münz, 2005) zeigt, erkennt man, wo man Keywords grundsätzlich unterbringen kann. Einerseits bietet sich der Kopf der HTML-Seite an, also der Bereich zwischen den Angaben `<head>` und `</head>`. Hier können Informationen über das Dokument, zum Beispiel Meta-Informationen wie Titel, Beschreibung, Schlagwörter, Name des Autors, etc. angegeben werden. Diese Angaben werden jedoch nicht im Browserfenster dargestellt.

Für die Suchmaschinen-Optimierung haben diese Angaben besondere Bedeutung, da Suchmaschinen die Wörter, die sich im Titel, also zwischen den Title-

Tags `<title>` und `</title>`, oder in der Meta-Angabe *description* befinden, besonders stark gewichten. Glöggler (2003) macht zusätzlich darauf aufmerksam, dass dem Titel und dem Meta-Tag *description* zusätzlich noch eine besondere Bedeutung zukommt, da diese Informationen als Beschreibung der Webseite in den Ergebnislisten der Suchmaschinen aufscheinen. Vergleichbar mit der Wichtigkeit des Anzeigentexts bei bezahlten Einschaltungen, bilden auch bei den klassischen Ergebnissen die Angaben, die in Titel und Beschreibung gemacht werden, die Entscheidungsgrundlage für den Benutzer, ob er die Seite besuchen wird. Andere Meta-Angaben, wie etwa *keywords*, werden von den meisten Suchmaschinen ignoriert, weil die Gefahr besteht, dass sie missbräuchlich verwendet werden.

Neben dem Kopf weist Abbildung 7 noch den Dokumentenkörper als zweites Element einer HTML-Datei aus. Zwischen den Body-Tags `<body>` und `</body>` werden alle Texte, Links, Bilder und andere multimediale Elemente angegeben, die bei einer Anzeige in einem Browser im Fenster sichtbar sein sollen. Für den Suchmaschinen-Optimierer bietet sich im Dokumentenkörper eine Vielfalt an Möglichkeiten um Keywords erfolgsversprechend zu platzieren. Glöggler (2003, 149) nennt in diesem Zusammenhang zum einen das *alt*-Attribut des `<img>`-Tags, das ursprünglich einen Hinweis enthielt, wenn ein Browser keine Darstellung von Bildern erlaubt hat. Zum anderen werden Kommentare als Möglichkeit vorgestellt, um Schlüsselwörter auf inhaltsarmen Seiten zu platzieren, ohne dass sie vom Benutzer gesehen werden.

Von noch zentralerer Bedeutung bei der Gestaltung des Dokumentenkörpers ist es allerdings, die Möglichkeiten, die HTML in Hinblick auf die Strukturierung des Textes vorsieht, voll auszunutzen. Suchmaschinen befinden etwa Begriffe, die eine besondere Textauszeichnung aufweisen, für besonders wichtig. Platziert man also Schlüsselbegriffe in einer Überschrift, also zum Beispiel zwischen `<h1>`- und `</h1>`-Tags, oder hebt man sie durch eine differenzierte Textauszeichnung, etwa Fettschreibung (`<b>...</b>` oder `<strong>...</strong>`) oder ähnlichem, hervor, bekommen die Keywords ein stärkeres Gewicht bei der Relevanzbewertung.

Neben dem Kriterium, wo die Deskriptoren vorkommen, wird errechnet, wie oft die Schlüsselbegriffe im Text verwendet wurden. Dabei gehen die Suchmaschinen wie klassische Information Retrieval Systeme vor: Es wird einerseits die relative Häufigkeit der Deskriptoren innerhalb des Dokuments errechnet,

Lewandowski (2005, 91) spricht in diesem Fall von dokumentspezifischer Wortgewichtung (WDF), andererseits wird die inverse Dokumentenhäufigkeit (IDF) errechnet, die Auskunft darüber gibt, wie häufig die Begriffe im gesamten Index vorkommen. Für die Suchmaschinen-Optimierung heißt das, dass ein Keyword umso wichtiger wird, je häufiger es im Dokument vorkommt und je seltener es im gesamten Index ist.

Letztlich ist auch die genaue **Position der Keywords** wichtig. So werden Begriffe, die am Beginn des Dokuments stehen stärker bewertet, als jene am Ende. Abhängig von der Anfrage, die der Benutzer an die Suchmaschine stellt, bekommen etwa bei Anfragen mit mehreren Begriffen zusätzliche Kriterien eine Bedeutung, wie der Abstand zwischen den Suchbegriffen, die Reihenfolge der Suchbegriffe in der Anfrage, Groß- und Kleinschreibung oder Sprache (Lewandowski, 2005, 91).

Auf diesem, oben beschriebenen, grundsätzlichen Muster der On-The-Page Methoden muss auch bei der Auswahl der Keywords Rücksicht genommen werden. Zwar gelten für die Erstellung der Keywordliste wiederum jene Angaben, die bereits in Kapitel 3.2.2 in Bezug auf Suchmaschinen-Werbung gemacht worden sind, allerdings mit der Einschränkung, sich je Web-Dokument auf wenige Begriffe zu konzentrieren. Glöggler (2003, 128) rät in diesem Zusammenhang zu drei bis fünf Keywords und zur „*thematischen Konsistenz*“ (Glöggler, 2003, 145), also dazu, thematisch nur ein Thema je Dokument zu erörtern.

3.3.2 Off-The-Page Methoden

Im Gegensatz zu On-The-Page Maßnahmen, haben Off-The-Page Maßnahmen das Ziel, die Suchmaschinen von der Qualität der Dokumente und Autorität der Urheberschaft zu überzeugen. Diese Qualitätsbewertung auf Basis der Off-The-Page Kriterien geschieht völlig unabhängig von der Anfrage oder dem Inhalt und dem Aufbau der Dokumente. Vielmehr stehen bei dieser Qualitätsbewertung die Besonderheiten des Internets, insbesondere von Hypermedia, im Mittelpunkt: Die Suchmaschinen analysieren zum Beispiel die Verzeichnisstruktur des Internetauftritts, die Änderungsfrequenz der Inhalte, wie gut eine Seite verlinkt ist und sogar wie häufig sie angeklickt wird. In Anlehnung an die theoretischen Grundlagen der Bibliometrie und der Zitationsanalyse entwickelten die Suchmaschinenbetreiber jene Verfahren zur Bewertung dieser

Kriterien, die heute typisch für Internet-Suchmaschinen als Web Information Retrieval Systeme sind.

Aktualität ist zum Beispiel ein Off-The-Page Faktor, auf den Suchmaschinen besonderen Wert legen. Eine häufige Änderungsfrequenz der Inhalte der Web-Dokumente hat einerseits zur Folge, dass die Suchmaschine die Inhalte als besonders aktuell und damit auch als qualitativ hochwertig einschätzt, andererseits hat es den Vorteil, dass die Website aufgrund der Änderungen in vergleichsweise kurzen Zeitabständen von den Webrobot-Systemen der Suchmaschine besucht wird und der Index regelmäßig aktualisiert wird.

Domainname und Verzeichnisstruktur werden als wichtiger Faktor für den Erfolg von Suchmaschinen-Optimierung besonders von Glöggler (2003, 169) betont. Die Meinung, dass sich die Domain-, Verzeichnis- und Dateinamen als Möglichkeit eignen, Keywords zu platzieren, wird jedoch nicht von allen Autoren facheinschlägiger Literatur geteilt: So bezweifelt etwa Thurow (2003, 84), dass sich Keywords in der URL entscheidend auf die Platzierung auswirken. Lewandowski (2005, 94) wiederum erwähnt in Bezug auf die Verzeichnisstruktur, dass die Suchmaschinen auch aus der Verzeichnisebene, in der das Dokument gespeichert wurde, auf die Wichtigkeit des Dokuments schließen, sodass ein Dokument eine höhere Bewertung bekommt, umso höher es in der Verzeichnishierarchie des Webserver abgespeichert ist. Darüber also, welche Bedeutung die URL und die Verzeichnisstruktur für die Suchmaschinen-Optimierung hat, ist man sich selbst in der Literatur nicht einig.

Keine widersprüchlichen Meinungen gibt es hingegen bei der Bedeutung der **Verlinkung innerhalb der Website und mit anderen Webseiten** im Internet. Hier wird aus der Linkstruktur des Internets nach einem Algorithmus ein Wert für die Qualität der Webseiten errechnet. Auf Basis des Science Citation Index (Lewandowski, 2005, 118) hat man im Laufe der Zeit eine Reihe von Algorithmen entwickelt, die neben der Anzahl der Verweise auch die Qualität der Verweise in die Kalkulation aufnimmt. Die grundlegende Überlegung dabei ist, dass ein Link nicht zufällig gesetzt wird, sondern Ausdruck der Wertschätzung für ein Web-Dokument und damit ein Qualitätsmerkmal darstellt. Als bekanntester dieser Algorithmen gilt PageRank des Marktführers Google. Eine genaue Beschreibung von PageRank und anderen Algorithmen zur Bestimmung der Link Popularität, wie etwa HITS oder Hilltop, findet man zum Beispiel bei Lewandowski (2005, 118-132). Durch die Komplexität dieser Algo-

rithmen ist es für die Suchmaschinen-Optimierung schwer geworden, auf den Faktor Link Popularität positiv Einfluss zu nehmen. Grundsätzlich gilt, dass sich wenige Links von qualitativ hochwertigen Seiten viel positiver auswirken, als viele Verweise von Seiten zweifelhafter Qualität (Thurrow, 2003, 113). Auf jeden Fall ist davon abzuraten, an dubiosen Linktausch-Programmen teilzunehmen. Sichergestellt sein muss jedoch immer, besonders vor dem Hintergrund der technischen Grundlagen der Suchmaschinen, dass die Webrobot Systeme der Suchmaschinen den Links problemlos folgen können, denn nur so kann gewährleistet werden, dass die verlinkten Dokumente auch gefunden und indiziert werden können. Um technischen Problemen (siehe Kapitel 3.3.3) zuvorzukommen, die den Crawling-Prozess der Suchmaschine erschweren, empfiehlt es sich, syntaktisch richtig programmierte HTML-Hyperlinks zu verwenden.

Neben Link Popularität, also Analysen der Linkstrukturen des Webs, beziehen die Suchmaschinen nutzungsstatistische Verfahren, die sogenannte **Klick Popularität** (click-through popularity), in die Qualitätsbewertung der Webseiten mit ein. Die Suchmaschine Direct Hit begann 1999 als erste Suchmaschine damit, die Tätigkeiten der Benutzer auf der Suchergebnisseite aufzuzeichnen und auszuwerten. Mittlerweile gibt es Direct Hit zwar nicht mehr, das Konzept Klick Popularität hat jedoch überlebt und wird heute von vielen Suchmaschinen als Teil der Bewertung von Internet-Dokumenten eingesetzt. Dabei wird gemessen, wie oft ein Link, der den Benutzer auf die Website bringt, angeklickt wurde, wie lange sich der Nutzer auf der Seite aufgehalten hat und wie oft er wieder zur Ergebnisseite zurückgekehrt ist (Thurrow, 2003, 117). Aufgrund zahlreicher Manipulationsversuche wird die Klick Popularität nicht mehr über die Suchmaschinenergebnisliste gemessen, sondern durch Toolbars (siehe Fußnote 21 auf Seite 39), die von beinahe jeder Suchmaschine angeboten werden und vom Internetbenutzer im Browser integriert werden (Lewandowski, 2005, 104).

3.3.3 Grenzen der Optimierung

Mit den bisher genannten Methoden kann man sehr gut die Platzierung einer Webseite beeinflussen. Bei der Optimierung von Webseiten gibt es allerdings auch einige Faktoren, die zwar in erster Linie keinen positiven Einfluss auf die Platzierung nehmen, aber zum Beispiel eine Aufnahme in den Datenbestand einer Suchmaschine erschweren oder verhindern können. Grundsätzlich kann

man diese Faktoren einerseits unter dem Titel technische Herausforderungen an eine Website zusammenfassen, andererseits setzen auch die Qualitätsanforderungen der Suchmaschinenbetreiber der Webseitenoptimierung Grenzen. Die wesentlichen Faktoren sind hier in weiterer Folge angeführt:

Technische Herausforderungen bestehen für den Webmaster typischerweise in der Konfiguration des Webserver oder in der Art und Weise, wie die Webseite aufgebaut ist. In Bezug auf die Serverkonfiguration stellt unter anderem eine schlechte Serververfügbarkeit ein großes Problem dar, denn die Webrobots der Suchmaschinen besuchen in regelmäßigen Abständen bereits indexierte Webseiten, um einerseits bei inhaltlichen Änderungen den Index zu aktualisieren, und andererseits Websites, die nicht mehr existieren, aus dem Index zu entfernen. Fällt der Webserver regelmäßig aus, schließt die Suchmaschine daraus, dass die Website dauerhaft entfernt wurde und löscht sie aus ihrem Index.

Was die Webseite selbst betrifft, die auf dem Webserver gespeichert wird, stellt alles außer syntaktisch richtig programmiertes HTML ein potentielles Problem für die Suchmaschinen und damit eine Herausforderung für den Suchmaschinen-Optimierer dar. Aber auch Framesets, also aus mehreren Einzelseiten zusammengesetzte Webseiten, können eine vollständige Erfassung aller Web-Dokumente einer Website erschweren (Glögler, 2003, 123). Zusätzlich zählen zu den Problemfeldern serverseitige dynamische Programmierung (ASP, PHP, JSP, ...) und clientseitige Scripts (JavaScript). Letztlich bedarf auch der Einsatz von Multimedia auf Websites einer besonderen Behandlung: Bilddateien, Videodateien und Flash-Animationen können von Suchmaschinen nur durch Meta-Informationen indexiert werden und Links, die von den multimedialen Inhalten auf andere Seiten verweisen, werden zum überwiegenden Großteil nicht verfolgt.

Zur Lösung des Problems bietet sich an, Dateien, die von der Suchmaschine nicht indexiert werden können oder sogar sollen, mit einer robots.txt-Datei unter Zuhilfenahme des Robots Exclusion Protocols von der Indexierung auszuschließen¹⁶. Stehen einer erfolgreichen Aufnahme in den Datenbestand der Suchmaschine die Hindernisse dynamischer Programmierung im Wege, kann man im Rahmen eines Paid Inclusion (bei weniger als 500 Einzelseiten) oder

¹⁶ Detaillierte Informationen bietet Glögler (2003, 166) und sind im WWW unter <http://www.robotstxt.org> zu finden.

Trusted Feed (bei mehr als 500 Einzelseiten) Programms einer Suchmaschine für die Indexierung bezahlen (Seda, 2004, 123). Diese Programme garantieren eine Aufnahme in den Index der Suchmaschine trotz technischer Hürden, haben jedoch keinen Einfluss auf die spätere Platzierung in der Suchergebnisliste.

Qualitätsrichtlinien der Suchmaschinen setzen vor allem unseriösen Anbietern von Suchmaschinen-Optimierung Grenzen. Denn neben den in Kapitel 3.3.1 und 3.3.2 genannten seriösen Methoden, die explizit von den Suchmaschinen befürwortet werden, gibt es eine Reihe von Tricks, die Lücken in den Bewertungsalgorithmen der Suchmaschinen ausnützen, um auch irrelevante Webseiten an die Spitze der Suchmaschinenergebnisse zu bringen. Zum überwiegenden Großteil nutzen diese Tricks die Schwäche der Suchmaschinen aus, kein JavaScript verarbeiten und Manipulationen mit Textauszeichnungen und -formatierungen erkennen zu können (Beispiele siehe unten). Man kann also mit relativ einfachen Mitteln Webseiten mit Texten und Links so aufbereiten, dass die Suchmaschinen die Webseite zu einem bestimmten Suchbegriff hoch bewerten, ohne dass einem menschlichen Seitenbesucher diese Manipulation auffällt. Die Suchmaschine indiziert also Inhalte, die für den Benutzer nicht sichtbar sind, was für den Benutzer der Suchmaschine zu einer massiven Verschlechterung der Qualität der Suchmaschinenergebnisse führt.

Für die Suchmaschinen stellt diese Form von Spam ein riesiges Problem dar: So schätzt etwa Wolfgang Sander-Beuermann (Geschäftsführer des Vereins SuMa e.V., zitiert in Hannemann/Jutzi/Körner (2005, 115)), dass ein Viertel aller Webseiten nur existieren, um die Ergebnisse von Suchmaschinen zu manipulieren. Und Lewandowski (2005, 56) erwähnt, dass die Suchmaschine Inktomi im Herbst 2003 von 5 Milliarden URLs, die sich zu diesem Zeitpunkt im Index ihrer Suchmaschine befanden, nur etwa 1,2 Milliarden für „indexierungswürdig“ hielt.

Als Gegenmaßnahmen stellen die Suchmaschinenbetreiber, wie zum Beispiel Google (2005a), Yahoo! (2005) und MSN (2006), auf ihren Hilfeseiten Informationen über ihre Qualitätsrichtlinien zur Verfügung und warnen die Webmaster davor, Schlupflöcher der Suchmaschinen auszunützen. Auch wenn im folgenden Abschnitt die wichtigsten Methoden, die gegen die Qualitätsrichtlinien der Suchmaschinen verstoßen, vorgestellt werden, sei der Leser vom Gebrauch dieser Mittel gewarnt: Viele dieser Tricks werden von den Suchmaschinen erkannt und rigoros mit einer schlechten Bewertung oder gar dem Ausschluss

aus dem Index bestraft. Eine ausführliche Beschreibung der gängigsten unerlaubten Manipulationsversuche finden sich bei Ott (2006).

Zusammengefasst zählen zu diesen unzulässigen Methoden unter anderem Texte und Links, die durch entsprechende Textauszeichnung (Schriftgröße, Schriftfarbe, ...) oder Positionierung im Quelltext (versteckte Ebenen, <noframes>-Bereich, <noscript>-Bereich, ...) den Algorithmus der Suchmaschine manipulieren. Als Extremform gilt das Vollstopfen mit Keywords, das sogenannte *keyword stuffing*, bzw. das *keyword stacking*. Eine weitere Form der Manipulation ist es, durch clientseitige Programmierung (*doorway pages*) oder serverseitige Mechanismen (*cloaking*) Suchmaschinen andere, extrem optimierte Inhalte zu übermitteln als einem menschlichen Besucher. Ebenso verstößt es gegen die Qualitätsrichtlinien der Suchmaschinen, doppelte Inhalte (Mirroring) anzubieten oder fremde Seiten zu kopieren (Pagejacking), um sich einen Vorteil zu verschaffen.

Haben die zuvor genannten Mittel das Ziel, die On-The-Page Faktoren zu manipulieren, zielen Link Farms darauf ab, die Link Popularität zu beeinflussen. Hier werden tausende Links ohne thematischen Zusammenhang auf einer Internet-Seite platziert. Mittlerweile funktioniert diese Methode nur mehr schlecht, da für die Kalkulation der Link Popularität nicht mehr nur die Anzahl der Links entscheidend ist, sondern vielmehr die Qualität.

Bei vielen der oben genannten Methoden scheint die Trennlinie zwischen Erlaubtem und Unerlaubtem teilweise nicht genau gezogen werden zu können. Manche Maßnahmen können derzeit von den Suchmaschinen nur schwer erkannt werden und vieles wird noch toleriert, weil es durchaus Anwendungsgebiete gibt, die den Einsatz von zum Beispiel Doorway Pages rechtfertigen. Allerdings ist die Wahrscheinlichkeit relativ hoch, dass sich die Situation schnell ändert: Die große Gefahr besteht zum einen durch die Weiterentwicklung der Algorithmen für die Identifizierung von Spam, um die Qualität der Ergebnisse zu verbessern. Zum anderen kann es durchaus sein, dass eine Webseite, zum Beispiel wegen eines Hinweises von einem Mitbewerber, aus dem Index genommen wird. Vor allem in Anbetracht des Schadens, der entstehen kann, wenn eine Webseite aus dem Index entfernt wurde, ist also das Risiko groß.

3.4 Überwachung und Erfolgskontrolle von Suchmaschinen-Marketing-Kampagnen

Sowohl für Suchmaschinen-Werbung, als auch für Suchmaschinen-Optimierung muss die Überwachung und Kontrolle des Erfolgs ein zentrales Element einer Kampagne darstellen. Denn schon während der Kampagne kann man durch ein regelmäßiges Monitoring Fehler beheben und gewisse Faktoren verbessern. Spätestens nach Abschluss der Kampagne sollte man auf Basis der gewonnenen Daten einen Eindruck von den Stärken und Schwächen der Kampagne gewinnen und nicht zuletzt erfahren, ob man die Ziele, die man sich bei der Planung der Kampagne gesetzt hat, auch wirklich erreicht hat. Diese Informationen sind außerdem der beste Ausgangspunkt für Verbesserungen.

Für die Erfolgskontrolle, insbesondere von Kampagnen, die das Ziel haben die Webseiten zu optimieren, nennt Glöggler (2003, 260) zum Beispiel das Website-Monitoring. Damit wird automatisiert sichergestellt, dass die Website ständig erreichbar ist und damit einer problemlosen Indexierung nichts im Wege steht. Stellt das System fest, dass die Website nicht mehr erreichbar ist, wird unverzüglich ein Verantwortlicher alarmiert, damit der Fehler behoben werden kann. Darüber hinaus ist es wichtig, die Webserver-Protokolle, die sogenannten Logfiles, auszuwerten. Aus diesen Serverprotokollen erhält man detaillierte Daten darüber, wieviele Anfragen zu welchen Seiten in einem gewissen Zeitraum an den Webserver gestellt wurden, welchem Link die Besucher gefolgt sind, um zur Website zu gelangen oder zum Beispiel welche Suchmaschine und Suchbegriffe sie verwendet haben. Diese Informationen stellen sich für gewöhnlich besonders bei der Weiterentwicklung der Website und für künftige Suchmaschinen-Marketing-Kampagnen als nützlich heraus. Schlussendlich muss auch die Rangposition einer Website regelmäßig überprüft werden, damit man schon während der Kampagne gewissen Veränderungen gegensteuern kann. Hierzu eignen sich sogenannte Rank Checking Programme, die es in einer Vielzahl von Ausführungen im Internet gibt.

Zusätzlich zu diesen grundlegenden Verfahren gibt es Analyseprogramme, die den speziellen Anforderungen von Suchmaschinen-Werbung entsprechen (Seda, 2004, 213). So werden von den Suchmaschinenbetreibern, aber auch von Drittanbietern Programme zur Verfügung gestellt, die es zum Beispiel erlauben, die sogenannte *click-through-rate* (CTR), also das Verhältnis zwischen *impressions* (Häufigkeit, wie oft eine bezahlte Einschaltung angezeigt wur-

de) und den *clicks* (Häufigkeit, wie oft eine Anzeige geklickt wurde) zu errechnen. Aus diesem Wert kann man ablesen, ob sich die Suchmaschinen-nutzer vom Text der Einschaltung angesprochen fühlen. Um den finanziellen Erfolg einer Suchmaschinen-Marketing-Kampagne in Form einer Return-On-Investment (ROI) Rechnung zu messen ist es hingegen wichtig, zu wissen, wie häufig ein Besucher das beworbene Produkt gekauft hat (*conversion rate*) und ob die Kosten für Suchmaschinen-Marketing diese Einnahmen nicht übersteigen.

4 Empirische Untersuchung von HTML-Quelltexten

Bisher wurde das Wesen des Marketinginstruments Suchmaschinen-Marketing erläutert. Die theoretischen Grundlagen bilden hierbei Informationen aus fach-einschlägiger Literatur. Mit einem Fokus auf Suchmaschinen-Optimierung wird im folgenden empirischen Teil der Arbeit untersucht, wie die Methoden von den Erstellern besonders erfolgreicher Webseiten angewendet werden. Hierzu werden ausgewählte Webseiten und deren HTML-Quelltexte qualitativ analysiert. Gefragt wird nach anfrageabhängigen Methoden (On-The-Page Methoden), insbesondere wie und in welchem Ausmaß sie angewendet worden sind. Ziel der Untersuchung ist es, herauszufinden, wie man eine Website auch ohne starke Anwendung von Optimierungsmethoden auf den ersten beiden Seiten der Ergebnislisten von Google, Yahoo! und MSN (siehe Abbildung 2 auf Seite 10) platzieren kann.

Aus den Ergebnissen der Untersuchung soll auf die Rolle der Methoden der Suchmaschinen-Optimierung für den Erfolg einer Webseite in Bezug auf ihre Platzierung in den Ergebnislisten der Suchmaschinen geschlossen werden. Dazu wird anhand der analysierten Webseiten hervorgehoben, welche Maßnahmen besonders erfolgsversprechend sind, wie diese Methoden eingesetzt werden müssen und insbesondere, welchen Einfluss die Qualität des Inhalts auf das Ranking der Suchmaschinen hat.

Als ein wichtiges Element der Suchmaschinen-Optimierung, werden die On-The-Page Kriterien der Webseiten für die Untersuchung herangezogen. Diese anfrageabhängigen Faktoren sind eindeutig und problemlos zu untersuchen, da sie aus der Programmierung des HTML-Quellcodes klar ersichtlich sind. Zusätzlich zu den On-The-Page Kriterien ziehen die Algorithmen der führenden Suchmaschinen für ihre Qualitäts- und Relevanzbewertung auch die sogenannten Off-The-Page Methoden heran. Besonders Kriterien, wie Verlinkung mit fremden Seiten (Link Popularität) oder Anzahl der Seitenbesuche (Klick Popularität), spielen hierbei eine wichtige Rolle. Empirisch sind diese Kriterien jedoch nicht analysierbar, da die Linkstruktur des gesamten Internet in einem vernünftigen Zeitaufwand nicht nachvollziehbar ist und genaue Daten bezüglich der Anzahl der Seitenbesucher nicht verfügbar sind. Im Rahmen der vorliegenden Untersuchung muss man sich daher auf die Angaben der Suchmaschinen

und Websitebetreiber oder auf Programme von Drittanbietern verlassen, die jedoch häufig nur wenige und kaum vertrauenswürdige Daten zurückliefern. Aufgrund dieser Schwierigkeiten, die aus der Komplexität der Suchmaschinenalgorithmen resultieren, liegt der Schwerpunkt der Untersuchung auf den On-The-Page Faktoren. Angaben zu den Off-The-Page Faktoren werden zwar gemacht, jedoch müssen sie aufgrund mangelnder Transparenz mit Vorsicht behandelt und kritisch hinterfragt werden, da sie auf Daten aus unzuverlässigen Quellen beruhen.

4.1 Untersuchungsgegenstand und Ausgangssituation

Als Grundlage für die empirischen Untersuchung dient eine Auswahl von Webseiten, die an einem Suchmaschinen-Optimierungswettbewerb (SEO-Contest) teilgenommen haben. Mit Hilfe von derartigen Wettbewerben wird versucht, die Arbeitsweise von Suchmaschinen besser kennen zu lernen, um Rückschlüsse auf suchmaschinenoptimiertes Webdesign ziehen zu können. Es werden Webmaster dazu aufgerufen, Webseiten so zu gestalten, dass sie zu einer vorher festgelegten Suchabfrage von ausgewählten Suchmaschinen möglichst hoch gerankt wiedergefunden werden. Um die Qualität der Suchmaschinenergebnisse außerhalb des Wettstreits möglichst wenig zu beeinträchtigen, werden von den Veranstaltern, häufig sind es einschlägige Websites oder Zeitschriften, Phantasiebegriffe als Schlüsselbegriffe ausgewählt. Beispiele für solche Schlüsselbegriffe von bereits vergangenen SEO-Contests sind die Wortschöpfungen *nigritude ultramarine* (2004), *Schnitzelmitkartoffelsalat* (2002) oder *Hommingberger Gepardenforelle*. Bis sie als Schlüsselbegriffe für einen SEO-Contest ausgewählt wurden, waren diese Begriffe und Begriffskombinationen jeder Suchmaschine gänzlich unbekannt. Heute liefern Abfragen mit diesen Suchphrasen zahlreiche, eigens für den Wettbewerb erstellte, suchmaschinenoptimierte Webseiten zurück.

Für die vorliegende Analyse wurde konkret der Wettbewerb um den Begriff *Hommingberger Gepardenforelle* untersucht. Dieser Wettbewerb wurde vom Heise Verlag in Hannover am 15. April 2005 ins Leben gerufen, um „*einen Einblick in die Rankingmechanismen der Suchdienste und aktuelle Trends der Optimierung - legitime wie unerwünschte*“ zu erhalten. An zwei Stichtagen, am 15. Mai 2005 und am 15. Dezember 2005, wurden unter den erfolgreichsten Webseiten, die zu den Suchbegriffen *Hommingberger Gepardenforelle* von den

Suchmaschinen Google, Yahoo, MSN und Seekport¹⁷ zurückgeliefert wurden, die Gewinner gekürt. Eine automatisierte quantitative Analyse der Methoden, die die Suchmaschinenoptimierer im Rahmen des Wettbewerbs verwendet haben, wurde in der Jännerausgabe der Zeitschrift c't veröffentlicht.

Suchmaschinen-Optimierungswettbewerbe locken viele erfahrene Webdesigner und sogenannte search engine optimisers (SEOs) an. Im Falle des Wettbewerbs um die Hommingberger Gepardenforelle haben eine Reihe von – hauptsächlich deutschsprachigen – Internetmarketing-Agenturen, aber auch Privatpersonen teilgenommen, um mit ihren Kenntnissen und Fähigkeiten in Bezug auf Suchmaschinen-Optimierung in den Wettbewerb mit anderen zu treten. Man kann also davon ausgehen, dass diese Webseiten nachvollziehbare Eigenschaften aufweisen, da sie sonst nicht im stark umkämpften Umfeld eines SEO-Contests bestehen könnten. Dabei können sich die teilnehmenden Webmaster voll auf die Methoden der Optimierung konzentrieren, da die Keywords vorgegeben sind, und damit die Planungsphase, wie sie in Kapitel 3 beschrieben wurde, wegfällt.

4.2 Vorgehensweise

Im Gegensatz zum Artikel des Heise Verlags, wird im Rahmen dieser Diplomarbeit der Schwerpunkt weniger auf die Methoden der Suchmaschinen-Optimierung selbst gelegt, sondern auf die Art und Weise, wie sie angewendet werden müssen, um langfristigen Erfolg in allen wichtigen Suchmaschinen zu erreichen. Dazu wurden am 15. Dezember 2005 um 11 Uhr, dem exakten zweiten Stichtag des SEO-Contest, die festgelegten Suchbegriffe *hommingberger gepardenforelle* in die österreichischen, deutschen und internationalen (US-amerikanischen) Suchinterfaces der Suchmaschinen Google, Yahoo! und MSN eingegeben und die ersten beiden Seiten der Ergebnisliste gespeichert. Im Laufe des selben Tages wurden die Websites, insbesondere die HTML-Quelltexte, aber auch robots.txt-Dateien, mit einem Browser-Plug-In¹⁸ lokal, für eine spätere Analyse gespeichert. Somit kann sichergestellt werden, dass die Website

¹⁷ In der vorliegenden Analyse wurde auf Daten der Suchmaschine Seekport verzichtet. Denn einerseits werden hier die gefundenen Webseiten nicht nur von einem Suchmaschinenalgorithmus, sondern auch von Menschen bewertet, und andererseits fällt der Suchmaschine außerhalb Deutschlands kaum eine Bedeutung zu.

¹⁸ Beim Browser-Plug-In handelt es sich um die SpiderZilla Erweiterung für den Firefox-Browser. Nähere Informationen findet man unter <http://spiderzilla.mozdev.org/>.

für die Untersuchung in genau jener Version zur Verfügung steht, die zum Stichtag des SEO-Wettbewerbs online war.

Für die Untersuchung stehen also 18 Screenshots von Suchergebnislisten¹⁹, die HTML-Quelltexte von mehr als 70 Webseiten und eine Reihe von Robots.txt-Dateien zur Verfügung. Zusätzlich wurden Informationen bezüglich Off-The-Page Kriterien, die vom Browser-Plug-In SEOpen²⁰ zurückgeliefert wurden, in einer Tabelle gespeichert. Obwohl diese Angaben direkt von den Suchmaschinen stammen, können sie nur als Vergleichswert hinzugezogen werden, da es sich nicht um die tatsächlichen Werte handelt. Je nach Quelle geht man davon aus, dass hier Schätzungen, Hochrechnungen oder einfach Zufallszahlen von den Suchdiensten zurückgegeben werden, um den Algorithmus für die Qualitätsbewertung, für den die Off-The-Page Methoden herangezogen werden, vor einer Manipulation zu schützen.

Aus den 70 Quelltexten wurden elf für eine detaillierte qualitative Analyse ausgewählt. Das Kriterium für die **Auswahl der Quelltexte** war eine Platzierung auf den ersten beiden Seiten der Ergebnislisten bei mindestens der Hälfte der betrachteten Länderinterfaces der Suchmaschinen Google, Yahoo! und MSN. Tabelle 2 zeigt die Platzierungen in den jeweiligen Länderinterfaces der Suchmaschinen und den Median aller Platzierungen.

Diese elf Webseiten sind in mehr als der Hälfte der Länderinterfaces der Suchmaschinen vertreten und zeichnen sich durch ihre gute Platzierung in allen Suchmaschinen aus. Sie weisen offenbar Eigenschaften auf, die für alle drei Suchmaschinen als Qualitätskriterium gelten. Um herauszufinden welche Eigenschaften das sind, wurden die HTML-Quelltexte, also die Programmierung der Webseiten, nach Methoden der Optimierung, insbesondere der On-The-Page Optimierung (siehe Kapitel 3.3.1), durchsucht und der inhaltliche Schwerpunkt der Webseite sowie die Intensität der vorgenommenen Optimierungsmaßnahmen beschrieben. Im folgenden Kapitel 4.3 werden die elf Webseiten und ihre Eigenschaften beschrieben. Auf Basis dieser gesammelten Daten wird schließlich in Kapitel 4.5 abgeleitet, welche Methoden von besonderer Bedeu-

¹⁹ Wie schon erwähnt, wurden für die Suchanfrage *hommingberger gepardenforelle* die ersten beiden Seiten der Ergebnislisten von google.at, google.de, google.com, yahoo.at, yahoo.de, yahoo.com, msn.at, msn.de und msn.com als Bilddatei gesichert.

²⁰ Nähere Informationen sind unter <http://seopen.com/firefox-extension/index.php> abrufbar.

URL	Google			Yahoo!			MSN			Median
	AT	DE	COM	AT	DE	COM	AT	DE	COM	
de.wikipedia.org/wiki/Hommingberger_Gepardenforelle	2	3	2	1	2	3	> 19	> 19	1	2
hommingberger-gepardenforelle.zielbewusst.de	8	9	8	2	1	4	> 19	17	> 20	8
www.danielgal.de/hommingberger-gepardenforelle/hommingberger-gepardenforelle.html	10	10	10	10	5	16	> 19	> 19	13	10
hommingberger-gepardenforelle.malkusch.de	11	11	11	14	9	> 20	3	2	> 20	11
www.hommingberger-gepardenforelle.de	1	1	1	> 20	> 20	> 20	11	9	> 20	11
www.kde.cs.uni-kassel.de/lehre/ss2005/googlespam/	9	8	9	13	8	20	> 19	> 19	14	13
www.heise.de/newsticker/meldung/58647	5	5	5	9	20	15	> 19	> 19	> 20	15
www.hommingberger-gepardenforelle.net	3	2	3	> 20	> 20	9	16	> 19	> 20	16
www.hommingberger-gepardenforelle-page.de	4	4	4	> 20	> 20	13	> 19	16	> 20	16
gepardenforelle.terrikay.de	18	19	18	12	7	18	> 19	> 19	> 20	18
www.hommingberger-gepardenforelle-blog.de	6	6	6	18	11	> 20	> 19	> 19	> 20	18

Tabelle 2: Platzierungen der ausgewählten Websites in den österreichischen, deutschen und internationalen Suchmaschineninterfaces von Google, Yahoo! und MSN.

tung für die Suchmaschinen-Optimierung sind und wie diese Methoden am aussichtsreichsten angewendet werden sollten.

Als **Untersuchungskriterien** sind im Rahmen der Analyse eine Reihe von Faktoren denkbar. Zuallererst wurde in den ausgewählten Quelltexten nach jenen *On-The-Page Methoden* gesucht, die in Kapitel 3.3.1 genannt wurden. Insbesondere wurde auf die Platzierung der Schlüsselbegriffe zwischen HTML-Elementen, die Häufigkeit des Auftretens der Suchphrase und die Position der Schlüsselbegriffe innerhalb des Quelltextes geachtet.

Weiters konnten auch die *Off-The-Page Methoden* (siehe Kapitel 3.3.2) bis zu einem gewissen Grad untersucht werden: Vor allem die Anzahl der Hyperlinks, die auf der indexierten Webseite gesetzt wurden, konnte festgestellt werden. Dabei wird zwischen Textankern, internen und externen Links unterschieden. Textanker zeigen auf eine andere Textstelle auf der selben Seite, von der auch der Link gesetzt wurde, interne Links verknüpfen einzelne Seiten innerhalb der selben Website und externe Links verweisen auf andere Websites. Gemeinsam mit den Angaben des Analyseprogramms SEOpen (siehe oben) bekommt man dadurch einen Eindruck über die Link Popularität der analysierten Webseiten. Sofern Informationen über die Verzeichnisstruktur und Aktualität der Webseite zur Verfügung standen, wurden auch diese in die Analyse miteinbezogen. Über die Klick Popularität der Seiten kann jedoch keine verlässliche Anga-

be gemacht werden. Hier bietet sich für eine Analyse lediglich der sogenannte Alexa Traffic Rank an, der jedoch wegen seiner recht zweifelhaften Ergebnisse nicht herangezogen wurde²¹.

Neben den On-The-Page und Off-The-Page Methoden ist es für die Zielsetzung der Untersuchung darüber hinaus wesentlich, die Intensität der Anwendung von Methoden der Suchmaschinen-Optimierung und die Qualität der Inhalte einzuschätzen. Schließlich soll herausgefunden werden, ob der langfristige Erfolg einer Webseite in Bezug auf Suchmaschinen-Optimierung von der intensiven Anwendung besonders raffinierter Optimierungsmethoden abhängt, oder ob die Einzigartigkeit des Inhalts auf lange Sicht über die Platzierung der Webseite in den Suchmaschinen entscheidet.

Um die *Intensität der Optimierungsmethoden* zu bewerten werden erneut die On-The-Page und Off-The-Page Methoden herangezogen. Maßgeblich für die Einschätzung der Intensität der Optimierungsmethoden ist es insbesondere, wie die Maßnahmen im Vergleich zu den anderen untersuchten Webseiten eingesetzt wurden. Je nachdem, wie sich die Anwendung der Optimierungsmaßnahmen darstellt, werden die Seiten anhand der folgenden, dreistufigen Skala zugeordnet:

1. Webseiten, die Optimierung *sehr intensiv und zielgerichtet* verwendet haben,
2. Webseiten, die SEO-Methoden *kaum oder zufällig* angewendet haben und
3. Webseiten, bei denen *keine Optimierung erkennbar* ist.

Um die *Qualität des Inhalts* der untersuchten Webseiten abzuleiten, bieten sich viele verschiedene Faktoren als Untersuchungskriterium an. Zum Beispiel könnte man untersuchen, wer die Ersteller der erfolgreich platzierten Webseiten sind: Privatpersonen, Multimediaagenturen oder andere Organisationen. Oder man stellt Vermutungen an, mit welcher Absicht die Seiten erstellt worden

²¹ Das Unternehmen Alexa gehört zur Unternehmensgruppe rund um Amazon. Mit Hilfe der Alexa Toolbar, einer optionalen Erweiterung für den Browser Microsoft Internet Explorer, werden Informationen zum Surfverhalten des Benutzers im Hintergrund aufgezeichnet und für eine Reihung der beliebtesten Websites (vgl. Klick Popularität) statistisch ausgewertet. Dadurch, dass es sich hierbei nicht um eine repräsentative Stichprobe handelt, ist die Aussagekraft des Alexa Rank jedoch nicht unumstritten.

sind, zum Beispiel als Wettbewerbsbeitrag oder als Berichterstattung über den Wettbewerb ohne Absicht, am Wettbewerb teilzunehmen.

Da jedoch diese Kriterien anhand der vorliegenden Informationen nur schwer zu überprüfen sind und damit die Aussagen über die Qualität des Inhalts auf Vermutungen basieren würde, wird im Rahmen der vorliegenden Untersuchung für die Bewertung der Qualität des Inhalts der Text selbst herangezogen. Wiederum anhand von drei Kategorien werden die dargestellten Inhalte bewertet:

1. Webseiten mit *objektiver Berichterstattung* über den Suchmaschinen-Optimierungswettbewerb,
2. Webseiten mit *Phantasietexten über die Hommingberger Gepardenforelle* und
3. Webseiten mit Texten *ohne Sinnzusammenhang*.

Auch diese Möglichkeit, die Qualität der Inhalte einzuschätzen, bringt gewisse Schwierigkeiten mit sich, weil sie keine eindeutige Bewertungsgrundlage darstellt. Die besondere Herausforderung ist es, die inhaltliche Qualität der Phantasietexte einzuschätzen: Die Texte ergeben zwar einen Sinn, ihr Nutzen für den Leser ist nicht mit einer objektiven Berichterstattung vergleichbar. In Bezug auf die Qualität liegen die Phantasietexte wohl zwischen den Texten mit objektiver Berichterstattung und den Texten ohne inhaltlichen Sinn.

4.3 Ergebnisse der qualitativen Analyse

Der folgende Abschnitt, der die Resultate der empirischen Untersuchung zusammenfasst, erfüllt im Wesentlichen drei Aufgaben: Er bildet die Grundlage für die Argumentation in Kapitel 4.5, veranschaulicht die Theorie aus den Kapiteln 2 und 3 mit Beispielen und Auszügen aus dem HTML-Code von erfolgreichen Webseiten und beschreibt, wie man Webseiten in Bezug auf Suchmaschinen-Marketing untersuchen kann, was abgesehen von einem Suchmaschinen-Optimierungswettbewerb zum Beispiel im Rahmen einer Analyse der Konkurrenz interessant ist.

Gemäß den Angaben der Suchmaschinen, wurden während des Suchmaschinen-Optimierungswettbewerb, der vom Heise Verlag am 15. April 2005 ins Leben

gerufen worden ist, bis zum 15. Dezember 2005 mehr als 60000 Seiten (Angabe der Suchmaschine MSN) zum Suchbegriff *hommingberger gepardenforelle* erstellt. Google will zu dieser Suchphrase sogar fast 3 Millionen Ergebnisse finden. Selbstverständlich gilt auch hier wiederum, dass man diesen Zahlen nicht allzu großes Vertrauen schenken darf. Bager (2005b) hat versucht die gesamte Ergebnisliste durchzuklicken, bekam insgesamt jedoch bei Google und Yahoo! nie mehr als 1000 und bei MSN nur bis zu 300 Ergebnisse präsentiert. Auch die Angaben zu den Backlinks, die man durch die Abfrage *link:hommingberger-gepardenforelle.de* erhält, kann man nur zu Vergleichszwecken heranziehen: Hier gibt etwa die Suchmaschine Google, so Bager (2005b), nur einen Bruchteil der tatsächlichen eingehenden Verweise aus, während Yahoo! und MSN tendenziell zu hohe Zahlen liefern.

Während des Wettbewerbs wurden eine Reihe von Methoden verwendet, die gegen die Qualitätsrichtlinien der Suchmaschinen verstoßen haben. Der Großteil dieser Seiten, die Lücken in den Algorithmen der Suchmaschinen ausnützen, wurden als Spam erkannt und mit einer verschlechterten Platzierung in den Ergebnislisten bestraft. Bager (2006) berichtet zum Beispiel von einer Seite, die sogenanntes URL Hijacking betrieben hat. Hier wird von einer schlecht bewerteten Seite serverseitig auf eine besser platzierte über den Statuscode 302 weitergeleitet, sodass die Suchmaschine zur Qualitätsbewertung der Ausgangsseite die Eigenschaften der Zielseite heranzieht.

Derart schwerwiegende Verletzungen der Qualitätsrichtlinien hatten bei den Suchmaschinen Google, Yahoo! und MSN nur kurzfristigen Erfolg und wurden schon bald aus den Ergebnislisten entfernt. Insbesondere MSN hatte jedoch auch mit geringfügigeren Manipulationen schwere Probleme: In den Ergebnislisten drei beobachteten Länderinterfaces belegte eine Multimediaagentur mit zweifelhaften Mitteln, wie Link Farms und Keyword Stuffing (siehe Abbildung 8), viele Plätze. Man spricht hier davon, dass die Ergebnislisten durch diese Manipulation verstopft worden sind. Außerdem fällt auf, dass die erste Seite der Suchmaschinenergebnisliste von MSN.at und MSN.de nur neun Resultate umfasst²². Wegen dieser Mängel scheint MSN für Bager (2006) mehr „*ein Spielfeld [...] um veraltete Spamming-Techniken auszuprobieren, denn ein ernst zu nehmender Suchhelfer*“ zu sein.

²² Es wird vermutet, dass hier ein Filter eingesetzt wurde, der die Seite von Wikipedia zur Suchphrase *Hommingberger Gepardenforelle* aus den ersten Ergebnislisten entfernt.

```

http://www.hommingberger-gepardenforelle-0.info/index.html
...
Hommingberger Gepardenforelle Hommingberger-Gepardenforelle Hommingberger
Gepardenforelle Hommingberger-Gepardenforelle Hommingberger Gepardenforelle
Hommingberger Gepardenforelle Hommingberger
...
<a href='_Hommingberger_Gepardenforelle_.html' target='_blank' ... >
Hommingberger Gepardenforelle</a>
...
Hommingberger Gepardenforelle Forelle Hommingberger Gepardenforelle Aufzucht
Forelle Impressum Sind die Hommingberger Gepardenforellen real? Diese Fische
sind eine relativ neue Schäpfung. Wenn man die ...
...

```

Abbildung 8: Beispiel für die unzulässige Methode des Keyword-Stuffing.

Am Stichtag (15. Dezember 2005) wies die Webseite der **Wikipedia**, erreichbar unter der URL http://de.wikipedia.org/wiki/Hommingberger_Gepardenforelle, die besten Platzierungen in allen betrachteten Suchmaschinenländerinterfaces (siehe Tabelle auf Seite 38) auf. Lediglich im österreichischen und deutschen Länderinterface der MSN Suche schaffte es die Webseite nicht auf eine der beiden ersten Suchergebnisseiten. Als Grund für diesen Misserfolg auf den beiden deutschsprachigen MSN-Interfaces vermutet man, dass hier ein Filter eingesetzt wurde, der die Seiten von Wikipedia wegen der wirtschaftlichen Konkurrenz zu einem kostenpflichtigen Produkt aus dem Hause Microsoft schlechter bewertet. Diese Annahme wird dadurch gestützt, dass im österreichischen und deutschen Länderinterface nur neun Resultate auf der ersten Seite der Ergebnisliste erscheinen. Bei anderen Suchbegriffen als *hommingberger gepardenforelle* werden – wie allgemein üblich – zehn Ergebnisse je Seite ausgegeben. Wie aus Tabelle auf Seite 38 ersichtlich ist, greift der Filter bei einer Suche auf MSN.com offensichtlich nicht; hier war die Wikipedia-Seite auf Platz eins.

Die Webseite selbst wurde als Artikel im Rahmen des deutschsprachigen Wikipedia Projekts erstellt, das das Ziel hat als freie Enzyklopädie seriöse Informationen zu unterschiedlichsten Themen für jedermann zugänglich zu machen. Der Domain- und Seiteninhaber ist die sogenannte Wikimedia Foundation mit Sitz in den Vereinigten Staaten von Amerika, die Inhalte werden jedoch von einer weltweiten Community erstellt und gewartet. Das heißt, dass jeder Besucher alle Artikel erstellen und auch bearbeiten kann. Am 15. Dezember 2005 bot die Webseite seriöse Informationen zum Suchmaschinenwettbewerb, auch wenn berichtet wurde, dass es mehrere Manipulationsversuche gegeben haben soll.

In Bezug auf On-The-Page Methoden wurden die Keywords im Kopf des HTML-Dokuments, nämlich im Meta-Tag *keywords* und im Title-Tag, positioniert (siehe Abbildung 9). Überraschenderweise wurde darauf verzichtet, die Suchphrase auch im Meta-Tag *description* zu verwenden.

```

Hommingberger_Gepardenforelle.html:
...
<head>
...
<meta name="KEYWORDS" content="Hommingberger Gepardenforelle" />
...
<title>Hommingberger Gepardenforelle - Wikipedia</title>
...
</head>
...

```

Abbildung 9: Auszug aus dem HTML-Quelltext der Wikipedia Seite als Beispiel für die Positionierung von Keywords im Kopf eines HTML-Dokuments.

Im Körper des Dokuments (`<body>...</body>`) wurde der Inhalt mit Ebenen, sogenannten `<div>`-Tags, organisiert²³. Die Phrase *Hommingberger Gepardenforelle* kommt nicht auffallend oft vor, laut Heise Online (2006) in einer Keyworddichte von nur 0,90 Prozent. Auch durch die Textauszeichnung ist keine intensive Verwendung von On-The-Page Methoden der Suchmaschinen-Optimierung erkennbar. Zwar wird die Suchphrase gleich zu Beginn in einer Überschrift erster Ordnung (`<h1>...</h1>`), etwas später zwischen `<b>` und `</b>` (Fettdruck) und ein weiteres Mal zwischen `<i>` und `</i>` (Kursivschrift) verwendet, es werden jedoch zahlreiche Möglichkeiten, etwa in weiteren Überschriften, alt-Texten von Bildern oder title-Attributen von Links, ausgelassen.

In Bezug auf die Off-The-Page Kriterien fällt auf, dass die Seite sehr viele ausgehende Hyperlinks aufweist. Von den 90 ausgehenden Links zeigen nur vier auf Webseiten, die nicht zu Wikipedia-Seiten zu zählen sind. Der Rest der Links sind interne Links, acht davon sind sogar Textanker und erfüllen damit die Aufgabe eines Inhaltsverzeichnis. Obwohl Wikipedia als eine große und wichtige Site im Internet gilt²⁴ weist der Artikel zur Hommingberger Gepardenforelle vergleichsweise wenige Backlinks (Google: 51, Yahoo!: 543, MSN: 1250) auf.

²³ Die Organisation in Layern hat den Vorteil, dass inhaltsarme Textteile am Ende der Seite platziert werden können.

²⁴ Die Suchmaschinen Google und Yahoo! haben mehrere Millionen von Seiten von Wikipedia im Index. Für Alexa, die über eine Toolbar die Bewegungen der Internetbenutzer auswerten, ist Wikipedia auf Rang 35 der wichtigsten Seiten im Internet.

Besonders bei Yahoo! und Google war die Seite **Zielbewusst**²⁵, die von Julian Grupp und Jens Diebold mit Unterstützung einer Multimediaagentur (its.diebold.de) erstellt wurde, erfolgreich. Lediglich bei MSN.at und MSN.com schien die Webseite nicht auf den ersten beiden Ergebnisseiten auf, obwohl sie durchaus sinnvollen Text enthält.

Der Dokumentenkopf enthält die Phrase *Hommingberger Gepardenforelle* im Title-Tag sowie im Meta-Tag *description* und *keywords*. Während im Title-Tag nur die Phrase vorkommt, finden sich in den beiden Meta-Tags jedoch neben der Phrase auch noch andere Wörter.

```
index.html:
...
<a href="index.html" title="Die Hommingberger Gepardenforelle">

</a>
...
```

Abbildung 10: Beispiel für die Möglichkeiten, wie Keywords in Links und Bildern untergebracht werden können.

Im Dokumentenkörper, der im Gegensatz zur Wikipedia-Seite mit einer Tabelle organisiert wird, kommt die Phrase ebenfalls gleich zu Beginn mehrere Male vor. Abbildung 10 zeigt, dass die Phrase im title-Attribut des Links, zwei mal im Pfad zum Bild und noch ein Mal im alt-Text des Bildes vorkommt. Danach wird die Phrase ein Mal zwischen `<h1>`-Tags platziert und mit der Textauszeichnung `<strong>...</strong>` vier Mal fett hervorgehoben. Abgesehen davon wurden keine Methoden der On-The-Page Optimierung im Zusammenhang mit der Suchphrase eingesetzt. Auch die Keyworddichte von 2,50 Prozent deutet darauf hin, dass bei der Erstellung der Seite die zur Verfügung stehenden SEO-Methoden nicht voll ausgeschöpft wurden.

Von den 23 ausgehenden Links, die auf der Webseite gezählt wurden, sind nur sechs externe Verweise. Der Großteil zeigt wiederum auf andere Seiten der Website, die jedoch häufig durch die externe robots.txt-Datei von der Indizierung ausgeschlossen sind. Auffallend hoch sind die Angaben, die von den Suchmaschinen in Bezug auf eingehende Links (Backlinks) gemacht werden: So will Google 124, Yahoo! 1360 und MSN 1430 Backlinks gefunden haben.

²⁵ URL: <http://hommingberger-gepardenforelle.zielbewusst.de>

Auch der Erfolg der Seite von **Daniel Gal**²⁶ konzentriert sich auf die Suchmaschinen Google und Yahoo!, während sie bei MSN – ebenso, wie die oben beschriebene Zielbewusst-Seite – nur beim internationalen Interface mit Platz 13 auf einer der ersten beiden Ergebnisseiten vorkommt.

Im Dokumentenkopf kommen, wie aus Abbildung 9 ersichtlich ist, die Suchbegriffe im Title-Tag und im Meta-Tag *keywords*, aber auch im Meta-Tag *description* vor. Außerdem erkennt man im HTML-Quelltext, dass die Inhalte des Dokumentenkörpers mit vier Ebenen, wiederum <div>-Layern, organisiert ist. Für den menschlichen Besucher sind jedoch nur drei dieser Ebenen sichtbar, denn eine Ebene, die besonders oft die Suchphrase und eine Reihe von Links enthält, ist speziell für die Suchmaschinen optimiert. Abbildung 11 zeigt, wie der Programmierer den suchmaschinenoptimierten Teil der Webseite versteckt hat: Über ein CSS (cascading style sheet), das Angaben zur Formatierung des Textes enthält, wurden zwei Ebenen exakt übereinander gelegt, wie die Positionsangaben top und left zeigen. Dabei liegt die Ebene *gepardenforelle-sage*, die durchaus sinnvollen Inhalt präsentiert, im Vordergrund (z-index:3), während die Ebene *gepardenforellen-links*, die den Blindtext und die versteckten Links enthält, dahinter (z-index:1) angeordnet ist. Grafische Browser²⁷ stellen also nur die Ebene *gepardenforelle-sage* dar, aber die Suchmaschine wertet auch die Ebene *gepardenforelle-links* aus, da sie die z-index-Angaben aus dem CSS ignoriert. Damit kann der Seitenbetreiber Texte und Links, die speziell für die Suchmaschine erstellt worden sind, vor dem menschlichen Besucher verstecken. Es wurde also offensichtlich ein beträchtlicher Aufwand für die Optimierung der Webseite betrieben.

Notwendig wurde dieser Verstoß gegen die Qualitätsrichtlinien der Suchmaschinen, da die Suchphrase zu häufig im sichtbaren Text vorkommt, nämlich zwölf Mal in 230 Wörtern, also in einer Häufigkeit von 5,22 Prozent. Eine derart hohe Keyworddichte ist für die meisten Suchmaschinen ein Indiz für Spam. Daher wurde in der versteckten Ebene sinnloser Text hinzugefügt, sodass die Keyworddichte auf 3,90 Prozent (Heise Online, 2006) reduziert werden konnte.

In Bezug auf die Textauszeichnung wurden im Gegensatz dazu nicht derart große Anstrengungen unternommen: Die Phrase wurde im sichtbaren Text le-

²⁶ URL: <http://www.danielgal.de/hommingberger-gepardenforelle/hommingberger-gepardenforelle.html>

²⁷ Die meisten Browser, wie Microsoft Internet Explorer, Mozilla Firefox, Opera, Safari oder Konqueror, sind grafische Browser.



Abbildung 11: Blindtext und versteckte Links auf den Seiten von Daniel Gal.

diglich jeweils ein Mal als Überschrift (<h1>-Tags), fett (-Tags) und kursiv (-Tag) angezeigt. Darüber hinaus wurde die Phrase zwei Mal jeweils im alt-Text und in einem title-Attribut von Bildern verwendet. Im unsichtbaren Bereich wurde die Phrase zusätzlich noch zwischen <h2>-Tags platziert.

Im gesamten Quelltext finden sich insgesamt 18 Links, wovon kein einziger auf eine externe Seite zeigt. Von den 18 Verweisen sind jedoch nur zwei für den Menschen sichtbar, während die übrigen 16 in der versteckten Ebene platziert sind. Zur Seite geben die Suchmaschine vergleichsweise wenige Backlinks zurück: Google registriert 26 eingehende Links, Yahoo! 343 und MSN 937. Besonders auffallend bei dieser Seite ist außerdem, dass sowohl der Verzeichnisname, als auch die Dateinamen die Suchwörter enthalten. Außerdem werden Dateien in unterschiedlichsten Formaten (.doc, .pps, .pdf, ...) zum Herunterladen angeboten.

Die Webseite, die von **Markus Malkusch**²⁸ erstellt worden ist, hat es geschafft mit relativ sinnvollen Texten in allen Suchmaschinen vertreten zu sein. Besonderen Erfolg hatte er bei MSN.at (Platz 3) und MSN.de (Platz 2), während er bei MSN.com und Yahoo.com nicht auf die ersten beiden Suchergebnisseiten kam.

Wie bei der Webseite von Wikipedia, wurde im Dokumentenkopf nur der Title-Tag und die Meta-Angabe *keywords* als Möglichkeit genutzt, um Suchbegriffe zu platzieren. Im Dokumentenkörper hingegen wurde die Suchphrase, aber auch die einzelnen Wörter *Hommingberger* und *Gepardenforelle* in vielen, teilweise auch in wenig verbreiteten Textauszeichnungen, wie `<kbd>`, `<abbr>` oder `<var>`, verwendet. Hervorstechend dabei ist die häufige Platzierung der Wörter in Überschriften (7 Mal `<h2>` und 9 Mal `<h3>`) sowie zwischen `<strong>`-Tags (14 Mal) und `<em>`-Tags (10 Mal). Der Aufbau mit `<div>`-Layern ermöglicht es zudem die Navigation am Ende des Quelltextes zu platzieren. Die Keyworddichte beträgt laut Heise Online (2006) 3,70 Prozent.

Die Webseite weist nur zwölf und wiederum ausschließlich interne Links auf. Auf die Möglichkeit, die Suchphrase in einem title-Attribut des Link-Tags unterzubringen wird verzichtet. Die Dateien, auf die verlinkt wurde, befinden sich jedoch stets in aussagekräftigen Unterordnern. Die Suchmaschinen haben 170 (Google), 1080 (Yahoo), beziehungsweise 1335 (MSN) Backlinks auf die Seite registriert.

Die offizielle **Wettbewerbsseite**²⁹ wurde vom Heise Verlag selbst ohne Rücksicht auf Optimierungsmethoden erstellt und hat fiktive Informationen über die Hommingberger Gepardenforelle als Inhalt. Obwohl sie seit 15. April 2005 nicht mehr aktualisiert wurde, ist sie in den beobachteten Länderinterfaces von Google mit drei ersten Plätzen höchst erfolgreich. Bei MSN.at und MSN.de erscheint sie immerhin auf den Plätzen elf und neun, bei MSN.com und den beobachteten Yahoo!-Interfaces kommt sie allerdings nicht unter die ersten 20 Ergebnisse.

Die Suchbegriffe sind in den Meta-Angaben *keywords* und *decription* sowie im Title-Tag untergebracht. Auch im Dokumentenkörper ist keine On-The-Page Optimierung erkennbar: Die Phrase wird gleich zu Beginn des Quelltextes ohne jegliche Textauszeichnung und später in einer `<h1>`-Überschrift positioniert.

²⁸ URL: <http://hommingberger-gepardenforelle.malkusch.de>

²⁹ URL: <http://hommingberger-gepardenforelle.de>

Darüber hinaus werden die Suchbegriffe in zwei `<h3>`-Überschriften, einem alt-Text eines Bildes sowie einmal fett, zwischen `<strong>`-Tags, und zwei Mal kursiv (`<i>`-Tags) geschrieben verwendet. Zwar ist die Platzierung der Phrase und der URL zur Seite in einem Kommentar (siehe Abbildung 12) auffällig und die Keyworddichte von 4,40 Prozent (Heise Online, 2006) vergleichsweise hoch, jedoch kann man bei der Wettbewerbsseite nicht von einer intensiven Anwendung von Optimierungsmethoden sprechen.

```
index.html:
...
<!-- Homepage der Hommingberger Gepardenforelle: http://www.hommingberger-gepardenforelle.de -->
...
```

Abbildung 12: Beispiel für die Möglichkeit Keywords in einem Kommentar zu platzieren.

Die Webseite enthält neun ausgehende Links, wovon alle auf interne Seiten verweisen und in keinem die Suchbegriffe im title-Attribut platziert werden. In Bezug auf die interne Verlinkung fällt besonders auf, dass man über die Navigation nicht mehr zur Startseite (`index.html`) zurückkehren kann, sondern nur auf die Datei `index_2.html`, die jedoch eine exakte Kopie der Startseite ist. Extern ist die Seite ausgesprochen gut verlinkt, was auch im Page Rank von 6 Niederschlag findet: Google zählt 114, Yahoo! 5720 und MSN 36082 Backlinks.

Im Rahmen einer Lehrveranstaltung hat die **Universität Kassel**³⁰ eine Seite als Beitrag zum Wettbewerb erstellt. Erklärtes Ziel, wie auch das Wort `google-spam` in der URL andeutet, ist die Manipulation der Suchmaschinenergebnisse. Man präsentiert hier einen besonders langen Text, in den die Suchphrase ohne Sinnzusammenhang eingefügt wurde, um die Schlagwortdichte zu erhöhen. Darüber hinaus wurden Studenten dazu aufgerufen, möglichst viele Links von anderen Seiten auf diesen Wettbewerbsbeitrag zu setzen. Dadurch haben es die Ersteller geschafft mit schwach optimiertem Text auch ohne sinnhaften Inhalt unter die ersten 20 Resultate der Ergebnisliste zu gelangen (Ausnahme: MSN.at und MSN.de).

Im `<head>` des HTML-Dokuments sind eine Reihe von Angaben zu finden. Hauptsächlich werden hier Angaben zur Universität und zur Lehrveranstaltung gemacht. Nur zwei Mal kommen die Suchworte *Hommingberger* und *Gepardenforelle* vor, nämlich im Title-Tag und im deutschsprachigen Keywords-Meta-Tag. Im Dokumentenkörper kommen die Suchbegriffe eher zufällig vor, da der

³⁰ URL: <http://www.kde.cs.uni-kassel.de/lehre/ss2005/googlespam/>

Großteil des Inhalts ein Text ist, in dem gewisse Wörter ohne Sinnzusammenhang mit der Suchphrase ersetzt worden sind. Die Suchbegriffe kommen zwar häufig in Überschriften vor, ansonsten jedoch in keiner besonderen Textauszeichnung. Hier wurde nur darauf geachtet, dass die Keyworddichte (laut Heise Online (2006): 2,60 Prozent) hoch genug ist, um in den Suchmaschinen als relevant bewertet zu werden.

Von der Webseite zeigen 43 Links weg. Davon verweisen alle außer vier auf andere Seiten der Universität Kassel. Darüber hinaus werden innerhalb des Dokuments sechs Textanker verwendet. Bei den Links wird das title-Attribut nicht verwendet, die Phrase kommt jedoch manchmal im Linktext vor. Eine Sonderstellung nimmt die Webseite auch bei den Backlinks ein: Auf der Webseite werden die Seminarteilnehmer aufgefordert, so viele Links wie möglich von anderen Seiten zu dieser zu setzen. Nichtsdestotrotz haben die Suchmaschinen Google (30 Backlinks), Yahoo! (1250 Backlinks) und MSN (142 Backlinks) nur vergleichsweise wenige eingehende Verweise gezählt.

Auch die Ankündigung des Suchmaschinen-Optimierungswettbewerb auf den Seiten der **Heise News**³¹ hat besonders gute Platzierungen in den Suchmaschinen Google und Yahoo! erreicht. Dass die Seite bei MSN keine Platzierung unter den ersten 20 Treffern erreicht, liegt eventuell daran, dass in der robots.txt-Datei der Robot namens MS Search 4.0 Robot von der Indexierung der Seite ausgeschlossen ist (siehe Abbildung 13).

```
robots.txt:  
...  
User-agent: MS Search 4.0 Robot  
Disallow: /  
...
```

Abbildung 13: Ausschluss eines Robots durch das Robots Exclusion Protocol in einer robots.txt-Datei.

Die Webseite stellt die erste Berichterstattung über den SEO-Contest vom 16. Mai 2005 dar und präsentiert einen kurzen, aber seriösen Inhalt ohne jeglichen Einsatz von Suchmaschinen-Optimierung. Im Dokumentenkopf ist die Phrase wiederum nur im Title-Tag und im Meta-Tag *keywords* zu finden. Der Meta-Tag *description* ist zwar vorhanden, enthält jedoch keine Suchbegriffe. Im <body> kommen die Suchbegriffe erst relativ spät und häufig ohne besondere Textauszeichnung vor. Lediglich ein Mal kommt die Phrase in einer

³¹ URL: <http://www.heise.de/newsticker/meldung/58647>

Überschrift zweiter Ordnung (<h2>-Tag) vor und auch die Keyworddichte ist mit 0,50 Prozent (Heise Online, 2006) besonders gering.

Für den Erfolg in den Suchmaschinen zeichnen wahrscheinlich die Off-The-Page Faktoren verantwortlich: Google zählt 89, Yahoo! 824 und MSN 1496 eingehende Links, was einen Google Page Rank von 6 zur Folge hat. Von der Website selbst zeigen 103 Links weg, wovon wiederum nur neun auf externe Seiten zeigen, unter anderem auf die offizielle Wettbewerbsseite (siehe oben).

Die Webseite der schweizer **Sitepoint** GmbH³² ist die einzige Seite im Rahmen der vorliegenden Analyse, die nicht von Deutschen erstellt worden ist. Mit durchaus sinnvollen und langen Texten schafften es die Betreiber insbesondere bei Google in die vordersten Bereiche der Suchmaschinenergebnislisten, während bei Yahoo! nur im internationalen Interface und bei MSN beim österreichischen Länderinterface ein Platz unter den ersten 20 Ergebnissen erreicht wurde.

Charakteristisch für die Seite ist die sorgfältige Erschließung der Seite im Dokumentenkopf. Wie Abbildung 14 zeigt, kommen die Suchworte hier in verschiedenen Variationen und Kombinationen vor. Ebenso viele Variationen und Kombinationen finden sich auch im Dokumentenkörper: Hier liegt der Schwerpunkt auf Überschriften (1 Mal <h1>- und 11 Mal <h2>-Tag), Fettschreibung (-Tag) und alt-Texte von Bildern (12 Mal). Die Keyworddichte liegt bei 2,50 Prozent (Heise Online, 2006).

Von der Seite gehen 27 Hyperlinks weg. Sechs davon zeigen auf externe Seiten, einer davon auf die oben besprochene Zielbewusst-Seite. Bis auf die Zielbewusst-Seite, die im Wettbewerb als Konkurrenz angesehen werden muss, weisen alle Links auch title-Attribute, wiederum mit den Suchbegriffen oder leichten Abwandlungen davon auf. Laut Google hat die Internet-Seite 150, laut Yahoo! 3030 und laut MSN 2370 Backlinks. Damit zählt die Sitepoint-Page unter allen betrachteten Seiten zwar nicht zu jenen Seiten, die die meisten eingehenden Links zu verzeichnen hat, jedoch scheint sie von besonders bedeutenden Seiten verlinkt zu sein, da sie einen Page Rank von 6 aufweist. Darüber hinaus fällt auf, dass besonders die Namen Bilddateien mit Keywords gespickt sind.

³² URL: <http://www.hommingberger-gepardenforelle.net>

```

index.html:
...
<head>
<title>Ihre Hommingberger Gepardenforelle.</title>
<meta http-equiv="Content-Type" content="text/html; charset=iso-8859-1" />
<meta http-equiv="content-language" content="de" />
<meta name="robots" content="index, follow" />
<meta name="title" content="Ihre Hommingberger Gepardenforelle." />
<meta name="Keywords" content="hommingberger gepardenforelle, hommingberger,
gepardenforelle, homming, berger, geparden, forelle, hommingberg, gepardenforellen" />
<meta name="Description" content="Alles über den Lebensraum der hommingberger
Gepardenforelle und das Wesen selbst." />
<meta name="abstract" content="Hommingberger Gepardenforelle - meist in den
Gebieten um Hommingberg anzutreffen ist hommingberger Gepardenforelle." />
<meta name="revisit-after" content="3 Days" />
<meta name="distribution" content="global" />
<meta name="rating" content="General" />
<meta name="audience" content="All" />
<link rel="stylesheet" type="text/css" media="screen" href="style.css" />
</head>
...

```

Abbildung 14: Ausführliche Erschließung eines Internet-Dokuments im Head einer HTML-Datei.

Ähnlich wie die Sitepoint-Seiten ist der Wettbewerbsbeitrag von **Tobias Mast**³³ im Dokumentenkopf erschlossen. Hier findet man die Suchbegriffe im Title-Tag sowie den Meta-Tags *title*, *description*, *abstract* und *keywords*. Im Dokumentenkörper kommen die Suchbegriffe mit einer Häufigkeit von 3,30 Prozent vor (Heise Online, 2006), allerdings kaum mit einer besonderen Textauszeichnung. Lediglich in einer <h1>-Überschrift, in einem fett geschriebenen Bereich (-Tags) und in einem title-Attribut eines Bildes kommt die Suchphrase vor. Trotzdem scheint der Text speziell für Suchmaschinen geschrieben zu sein, wie Abbildung 15 zeigt: Damit die Suchmaschine die Schlagwörter in der richtigen Schreibweise indexiert, werden Apostrophe und Bindestriche verwendet.

```

index.html:
...
<b>Dies ist meine Hommingberger Gepardenforelle 'n Seite.</b>
...
- Restaurant zum Hommingberger-See-Blick<br>
Adresse:<br>
Tolle-Gepardenforelle 'n-Straße 133
12039 Hommingbergerhausen<br>
Tel.: 01113/777111<br>
...

```

Abbildung 15: Beispiel für einen optimierten Text.

Insgesamt gehen von der indexierten Seite 20 Hyperlinks weg. Nur vier davon zeigen auf externe Seiten. Bei internen Links und bei Bildern kommt häufig ein Schlagwort auch im Dateinamen vor. Darüber hinaus fällt in Be-

³³ URL: <http://www.hommingberger-gepardenforelle-page.de>

zug auf die Off-The-Page Kriterien auf, dass die Suchmaschinen eine enorme Menge an Backlinks gefunden haben: Google gibt für die Seite 537, Yahoo! 5820 und MSN 1490 eingehende Verweise zurück. Innerhalb der Site kommt man jedoch, ebenso wie auf der offiziellen Wettbewerbsseite nicht mehr auf die `index.html` zurück, sondern nur auf eine Kopie der Seite mit dem Dateinamen `index_2.html`.

Eine der wenigen Frauen, die am Suchmaschinen-Optimierungswettbewerb teilgenommen haben, ist **Teresia Kuhr** mit ihrer Seite, die unter der URL `http://gepardenforelle.terrikay.de` im Internet abrufbar ist. Allerdings konnte sie ihre Seite nur bei Google und Yahoo! unter die Top-20 Ergebnisse bringen, während sie es bei den Länderinterfaces von MSN nicht schaffte auf die ersten beiden Seiten der Ergebnisliste zu gelangen. Auch sie versucht mehr oder weniger sinnvolle Inhalte über den fiktiven Fisch auf der Seite zu präsentieren, teilweise jedoch durch intensive Anwendung von Methoden der Suchmaschinen-Optimierung.

Beim ersten Blick auf die Webseite fallen dem Besucher die Codierungsfehler bei Umlauten und Sonderzeichen auf. Außerdem fehlen im Quelltext `<head>`- und `<body>`-Tags³⁴. Weiters verdienen die Meta-Tags Aufmerksamkeit: Hier wurde die Suchphrase einfach in möglichst vielen Meta-Angaben, wie *keywords*, *description*, *author*, *copyright*, *page-type* oder *page-topic*, angeführt. Im Dokumentenkörper wurden die Schlagwörter jeweils ein Mal in den Überschriften `<h1>`, `<h2>` und `<h3>` verwendet. Häufiger fand die Textauszeichnung `<strong>` als Platz für die Keywords Anwendung. Jedes der beiden Bilder wurden mit alt-Texten und title-Attributen, das die Suchworte beinhaltet, versehen, sodass die Keyworddichte mit 4,80 Prozent (Heise Online, 2006) vergleichsweise hoch ausfiel.

Von den 19 Links, die die indexierte Webseite beinhaltet, zeigt nur einer auf eine Seite ausserhalb der Wettbewerbsseite. Die Seiten, auf die die internen Links zeigen haben zumeist die Suchbegriffe im Dateinamen. Bei der internen Verlinkung der Site wurde wiederum – wie bei der offiziellen Wettbewerbsseite und den Seiten von Tobias Mast – darauf geachtet, dass man über die Navigation nicht auf die Startseite zurückkehren kann, sondern nur auf eine Kopie mit dem Dateinamen `index_2.html`.

³⁴ Bei einer Überprüfung der Validität der Programmierung auf `http://validator.w3.org` wird trotz der Auslassungen zurückgegeben, dass die Seite dem Standard *HTML 4.01 Strict* entspricht.

Die Multimediaagentur **COM.POSITUM** schaffte es mit einem Weblog, das unter der Adresse <http://www.hommingberger-gepardenforelle-blog.de> abrufbar ist und differenzierte Informationen anbot, bei Google auf die erste Suchergebnisseite und bei Yahoo.at und Yahoo.de immerhin auf die zweite Seite der Ergebnisse, während ihnen dieser Erfolg bei Yahoo.com und den beobachteten Suchinterfaces von MSN verwehrt blieb.

Im Dokumentenkopf wurden der Title-Tag sowie die Meta-Angaben *keywords* und *description* mit den Suchbegriffen *Hommingberger Gepardenforelle* gefüllt. Der Dokumentenkörper enthält die Suchbegriffe schon zu Beginn und acht Mal in einer der 11 <h1>-Überschriften, sowie ein Mal in einer der 4 <h3>-Überschriften. Ansonsten wurde noch drei Mal die Fettschreibung mit -Tags angewendet.

Auf der Webseite finden sich 97 ausgehende Links, wovon zwei eMail-Links und 19 externe Verweise sind. Durch ihren attraktiven Inhalt scheint die Seite auch viele Backlinks gewonnen zu haben, wie Google mit 408, Yahoo! mit 6150 und MSN 3951 angeben. Außerdem fällt auf, dass die Seite Dateien in vielerlei Formaten (zum Beispiel PDFs sowie Bild-, Audio- und Videodateien) und die Suchergebnisseiten zur Suchphrase *hommingberger gepardenforelle* verlinkt. Nur wenige interne Links und die Verweise zu den Ergebnisseiten der Suchmaschinen weisen ein title-Attribut auf.

4.4 Zusammenfassung

Nach der Beschreibung der Ergebnisse, die aus der Analyse der Webseite hervorgingen, werden nun die angewendeten Methoden zusammengefasst dargestellt. Zuerst wird, unterteilt nach On-The-Page und Off-The-Page Maßnahmen, besprochen, welche Methoden verwendet wurden. Da diese Kriterien von den elf erfolgreichen Webseiten verwendet wurden, kann man daraus schließen, dass sie die Platzierung einer Webseite in den Ergebnislisten positiv beeinflussen. Danach wird auf die Intensität der Methoden der Suchmaschinen-Optimierung und die Qualität der Inhalte aller Seiten eingegangen. Diese Ausführungen bilden den Ausgangspunkt für die Schlussfolgerungen in Kapitel 4.5.

4.4.1 On-The-Page Kriterien

Aus der Verwendung der On-The-Page Kriterien lassen sich aus den Ergebnissen gewisse Schwerpunkte herauslesen. Besonders im Kopf der untersuchten HTML-Quelltexte scheint die **Platzierung der Schlüsselbegriffe** im `<title>`-Tag und im Meta-Tag `keywords` wichtig zu sein. Während die Bedeutung des Title-Tags auch in der Literatur betont wird, überrascht die weit verbreitete Verwendung des Meta-Tags `keywords`. Angeblich wird diese Meta-Angabe von den Suchmaschinen ignoriert, da es in der Vergangenheit immer wieder für Manipulationen, wie zum Beispiel `keyword stuffing` (siehe Abbildung 8) missbraucht wurde. Mit einer derart häufigen Verwendung hätte man eher beim Meta-Tags `description` rechnen müssen. Die Meta-Angabe `description` wurde jedoch ebenso, wie die Angabe `abstract`, nur von vergleichsweise wenigen untersuchten Webseiten angewendet. Die Platzierung von Schlagwörtern in allen übrigen Angaben, wie `author`, `copyright`, `page-type` und `page-topic`, zählen eher zur Ausnahme als zur Regel (siehe Tabelle auf den Seiten 60 und 61).

Im Dokumentenkörper werden die Suchbegriffe einerseits in Überschriften, dabei vorzugsweise zwischen `<h1>`-Tags, und andererseits in fett gedruckten Bereichen platziert. Für den Fettdruck wird in den elf ausgewählten Webseiten häufiger die Textauszeichnung `<strong>` verwendet als das `<b>`-Tag. Auch wenn für den Betrachter des Texts kein Unterschied zwischen den beiden Formen erkennbar ist, scheinen sich insbesondere die stark optimierten Seiten von der Verwendung des `<strong>`-Tags einen positiven Effekt zu erwarten. Ähnlich breite Anwendung findet die Platzierung von Schlüsselwörtern in alt-Texten von Bildern und `title`-Attributen von Links, wie sie zum Beispiel in Abbildung 10 dargestellt ist. Auf den Webseiten von Daniel Gal, Tobias Mast und Teresia Kuhr wird das `title`-Attribut auch bei Bildern verwendet³⁵. Wiederum eine Ausnahme ist überraschenderweise die kursive Textauszeichnung mit `<i>`-Elementen, sowie alle übrigen verwendeten Möglichkeiten der Text hervorhebung.

In Bezug auf die **Schlagwortdichte** bewegen sich die Seiten zwischen 0,5 Prozent und 4,8 Prozent. Ein Einfluss der Keyworddichte auf die Platzierung in

³⁵ Münz (2005) zählt das `title`-Attribut zu den Universalattributen. Das heißt, dass sie im Großteil aller HTML-Elemente verwendet werden dürfen (<http://de.selfhtml.org/html/referenz/attribute.htm>). Über eine positive Wirkung auf die Platzierung in Suchmaschinen ist aus der verwendeten Literatur nichts zu erfahren.

den drei Suchmaschinen Google, Yahoo! und MSN ist auf Basis der vorliegenden Untersuchung vordergründig nicht erkennbar. Denn die erfolgreichste Seite (Wikipedia) weist mit 0,9 Prozent die zweitgeringste Suchworthäufigkeit auf. Auffallend ist jedoch, dass besonders die Seiten, die von Privatpersonen erstellt worden sind, eine besonders hohe Schlagwortdichte aufweisen. Möglicherweise musste hier ein Mangel an Popularität, im Sinne der Off-The-Page Methoden, mit einer häufigen Verwendung der Suchphrase wett gemacht werden. Für die Multimediaagenturen, die häufig ein bemerkenswertes Netz von Webseiten für den Linktausch betreiben, scheint es in Bezug auf die Keyworddichte wichtiger zu sein, von den Suchmaschinen nicht als Spam erkannt zu werden. Google und Yahoo! bestrafen die Webseiten insbesondere bei leicht erkennbaren Spamming-Methoden, wie keyword stuffing, mit einer schlechten Bewertung, die sich theoretisch auf das gesamte Linknetzwerk der Agentur negativ auswirken kann. Unter den betrachteten Suchmaschinen bildet MSN die Ausnahme bei der Erkennung von Spam: Wie oben bereits erwähnt, berichtet Heise Online (2006), dass auf der ersten Ergebnisseite der MSN-Suche zwei Webseiten einer Agentur mit einer Keyworddichte von über 10 Prozent zu finden waren. Außerdem verlinkten sich die Seiten fast ausschließlich untereinander, was von Suchmaschinen von vergleichbaren Format normalerweise problemlos als Spam erkannt wird (Bager, 2006).

Auch bei der **Positionierung der Suchbegriffe** im Text wurde beinahe ausschließlich darauf geachtet, dass die Suchphrase *Hommingberger Gepardenfelle* möglichst ganz am Beginn des dargestellten Textes vorkommt, wodurch sie für die Suchmaschinen zusätzliche Bedeutung bekommen. Die einzige Ausnahme bilden die Heise News Seite und der Wettbewerbsbeitrag der Universität Kassel: Hier kommt die Suchphrase erst relativ spät, teilweise erst nach mehr als 100 HTML-Code-Zeilen vor. Wichtig in diesem Zusammenhang ist nicht zuletzt der Aufbau der Webseite: Um das Erscheinungsbild der Webseite für einen menschlichen Besucher ansprechend zu gestalten, bieten sich im Wesentlichen drei Möglichkeiten der HTML-Programmierung an, nämlich Framesets, Tabellen und Ebenen³⁶. Unter den Webseiten, die im Rahmen der empirischen Untersuchung betrachtet wurden, hat der Großteil der Seitenersteller Tabellen für das Layout verwendet. Ebenen, die sich insbesondere in

³⁶ Nähere Informationen zu den Vor- und Nachteilen der Möglichkeiten, die HTML für das Layout einer Webseite bietet, finden sich für Framesets unter <http://de.selfhtml.org/html/frames/layouts.htm>, für Tabellen unter <http://de.selfhtml.org/html/tabellen/layouts.htm> und für Ebenen unter <http://de.selfhtml.org/html/text/bereiche.htm>.

Kombination mit cascading style sheets (CSS) besonders flexibel positionieren lassen und es damit ermöglichen, inhaltstragende Bereiche an den Beginn des Quelltextes zu stellen, wurden nur von vier untersuchten Dokumenten verwendet. Framesets hingegen wurden von keiner der untersuchten Webseiten verwendet. Sie scheinen eine zu große Herausforderung (siehe Seite 29) zu sein, da sie besonders in Bezug auf Suchmaschinen-Optimierung zu viele Probleme verursachen und kaum Vorteile gegenüber den anderen Möglichkeiten bieten.

4.4.2 Off-The-Page Kriterien

Da viele Off-The-Page Kriterien, die in Kapitel 3.3.2 theoretisch erläutert sind, in ihrem praktischen Einsatz oft nicht eindeutig nachvollzogen werden können, beschränkt sich die Zusammenfassung der Ergebnisse in Bezug auf die Off-The-Page Methoden auf Kriterien, wie URL, Datei- und Verzeichnisnamen und die Links, die auf den indexierten Webseiten gesetzt wurden. Auf Fragen der Aktualität der Inhalte wird nur dann eingegangen, wenn Daten von den Suchmaschinen, den Webseitenerstellern oder den Webservern vorliegen. Das Kriterium Backlinks muss mit einer gewissen Skepsis behandelt werden, da die Zahlen, die von den Suchmaschinen zurückgegeben werden, wenn überhaupt, dann nur zum Zwecke eines Vergleichs herangezogen werden dürfen. Überhaupt keine aussagekräftigen Daten liegen zur Klick Popularität der untersuchten Seiten vor. Lediglich der Traffic Rank von Alexa, der Webseitenbesuche über ein Browser Plug-In – die sogenannte Alexa Toolbar – aufzeichnet und statistisch auswertet, gibt einen Eindruck von der Klick Popularität einzelner Domänen, jedoch nicht von einzelnen Webseiten.

Beim überwiegenden Teil aller Webseiten, die am SEO-Wettbewerb teilgenommen haben, finden sich beide **Suchbegriffe in der URL**. Auch wenn über den Effekt auf die Platzierung in den Ergebnislisten keine Einigkeit unter den Spezialisten herrscht, hat kaum ein Websitebetreiber darauf verzichtet, wenigstens eines der beiden Bestandteile der Phrase auch in der URL zu platzieren. Unter den elf analysierten Webseiten finden sich nur zwei URLs, die gänzlich auf die Erwähnung eines Keywords verzichten, nämlich die Seite der Universität Kassel und jene von Heise News. Der Wettbewerbsbeitrag von Teresia Kuhr wiederum führt nur das Wort *gepardenforelle* in seiner URL auf. Alle übrigen untersuchten Webseiten führen die Phrase mindestens ein Mal in der URL an, jene von Daniel Gal hat enthält die Phrase sogar zwei Mal. Die betrachteten

Webseiten weisen zum Großteil die deutsche Top-Level-Domain (TLD) .de auf und sind zum überwiegenden Teil in Deutschland gehostet. Die Ausnahmen sind hierbei Wikipedia und Sitepoint, die ihre Seiten unter den länderübergreifenden TLDs .org, beziehungsweise .net registriert haben und auf einem US-amerikanischen (Wikipedia), beziehungsweise schweizer Server (Sitepoint) gespeichert haben. Um die Vermutung bestätigen zu können, dass die Top-Level-Domain oder das Land, in dem eine Seite gehostet ist, die Platzierung in bestimmten Länderinterfaces beeinflussen, liegen in der vorliegenden Analyse zu wenige Beweise vor.

Auch die **Namen der Dateien und Verzeichnisse**, auf die von den untersuchten Seiten verlinkt wurde, können die vorgegebenen Suchbegriffe enthalten. Beispiele hierfür finden sich in den Quelltexten der Webseiten der Agenturen Sitepoint und COM.POSITUM, aber auch auf der Seite von Markus Malkusch. Hier wurden die Namen von HTML-Dateien sowie von Texten, Bildern, Videos und Musikstücken in unterschiedlichsten Formaten mit den Schlüsselbegriffen versehen, in der Hoffnung auf einen positiven Effekt auf die Platzierung in den Suchmaschinenergebnislisten.

Das Off-The-Page Kriterium **Aktualität**³⁷ kann nur schwer verglichen werden, da hier Informationen aus verschiedensten Quellen zur Verfügung stehen. Zum einen kann man sich auf die Angaben der Webseitenbetreiber verlassen, sofern welche vorhanden sind, zum anderen übergibt der Webserver – je nach Konfiguration – das Datum der letzten Änderung³⁸ und nicht zuletzt speichert die Suchmaschine das Datum, wann eine Kopie der Seite im sogenannten Cache archiviert hat. Basierend auf den vorliegenden Daten erkennt man die Tendenz, dass beinahe alle untersuchten Webseiten versucht haben, ihre Seite häufig zu aktualisieren. So wurden in der Woche vor dem 15. Dezember auf dem Weblog von COM.POSITUM, dem Wikipedia-Beitrag sowie auf den Seiten von Daniel Gal, Markus Malkusch und der Universität Kassel Änderungen festgestellt. Ende November und Anfang Dezember wurden die Seiten von Sitepoint, Zielbewusst und Tobias Mast aktualisiert, während die Webseite von Teresia Kuhr zuletzt am 25. Juli 2005 (Last Modified) verändert wurde. Die beiden

³⁷ Um das Off-The-Page Kriterium Aktualität (siehe auch Seite 27) zu beeinflussen, ist es nicht unbedingt notwendig, den Inhalt der Seite zu ändern, sondern es reicht oft schon eine Änderung im HTML-Quelltext, die sich auf das Erscheinungsbild oder den Inhalt der Webseite nicht auswirkt.

³⁸ Webserver können so eingestellt werden, dass sie das Datum der letzten Änderung an jeden Besucher im Header des HTTP-Protokolls mitschicken. Dieser Server Header namens Last Modified wird laut Glöggler (2003) auch von den Suchmaschinen ausgelesen.

Seiten Heise News und die offizielle Wettbewerbsseite, die vom Heise Verlag ins Leben gerufen wurden, sind beide laut Bager (2006) niemals aktualisiert worden.

Das zentrale Element der Off-The-Page Kriterien eines Suchmaschinenalgorithmus ergibt sich aus der Analyse der **Linkstruktur**. Für die vorliegende empirische Untersuchung stützen sich die Ausführungen hauptsächlich auf die Verweise, die von der betrachteten Webseiten weg zeigen. Diese sind anhand der gespeicherten HTML-Quelltexte genau nachvollziehbar und analysierbar. Für die Angaben bezüglich der Links, die von fremden Seiten auf die beobachtete Webseite zeigen, die sogenannten Backlinks, wurde im Rahmen der vorliegenden Analyse auf Daten zurückgegriffen, die die Suchmaschinen über das normale Suchinterface zurückgeben und als nicht vertrauenswürdig gelten.

Was die Anzahl jener Links betrifft, die auf den analysierten Webseiten zu finden waren und von dort weg zeigen, so weisen nur wenige Seiten sehr viele Links auf. Dazu zählen die Wikipedia Seite (90 Links), die Heise News Seite (103 Links) und die COM.POSITUM Seite (97 Links). Auch die Webseite der Universität Kassel verweist mit 43 Links auf viele Seiten. Alle übrigen Seiten haben weniger als 27 Hyperlinks, wovon auf der offiziellen Wettbewerbsseite mit neun Verweisen am wenigsten zu zählen waren. Typischerweise zeigen von diesen Links vier bis neun auf fremde Seiten und sind damit externe Verweise. Besonders wenige externe Verweise haben die Seite von Teresia Kuhr sowie die Seiten von Daniel Gal, Markus Malkusch und die offizielle Wettbewerbsseite, während das Weblog von COM.POSITUM vergleichsweise viele externe Links, nämlich 19 aufweist. Man erkennt also, dass von den untersuchten Webseiten viele interne Links weg zeigen. Bei drei betrachteten Websites erkennt man zusätzlich, dass es möglichst vermieden wurde, dass von anderen internen Seiten Links zurück zeigen. Insbesondere auf den Seiten von Tobias Mast, Teresia Kuhr und der offiziellen Wettbewerbsseite kann man über die normale Verlinkung nicht mehr auf die Startseite (`index.html`) zurückkehren, sondern nur auf eine exakte Kopie der Startseite, die mit dem Dateinamen `index_2.html` gespeichert wurde. Ob es sich hierbei um eine Optimierungsmethode oder um eine zufällige Gemeinsamkeit handelt, kann anhand der vorliegenden Informationen abschließend nicht geklärt werden.

Die vorliegenden Daten bezüglich der Backlinks seien hier nur kurz angerissen, da sie für die Schlussfolgerungen nur eine geringe Rolle spielen: Laut

Google weisen die Seiten von Tobias Mast (537 Backlinks) und die Webseite der Multimediaagentur COM.POSITUM (408 Backlinks) mit Abstand die meisten Backlinks auf. Auch für Yahoo! weisen diese beiden Seiten sehr viele Backlinks auf, nämlich 6150 (COM.POSITUM) und 5820 (Tobias Mast). Zusätzlich scheint auch die offizielle Wettbewerbsseite viele Links, nämlich 5720, von fremden Seiten bekommen zu haben. Im Index von MSN hat die offizielle Wettbewerbsseite die meisten Backlinks. Allerdings ist es offensichtlich, dass die Zahl von 36082 viel zu hoch und absolut unrealistisch ist. Mit großem Abstand folgt wiederum das Blog von COM.POSITUM (3951 Backlinks). Zieht man den Google Page Rank, der die Anzahl und Qualität der Linkstruktur in einer Zahl ausdrückt, für eine Einschätzung der Link Popularität heran, sind es die offizielle Wettbewerbsseite und die Heise News Seite sowie die Seite von Sitepoint, die sich mit ihrem Page Rank von 6 von den übrigen Webseiten (Page Rank 5) abheben.

4.4.3 Intensität der Optimierung und inhaltliche Qualität

Hauptaugenmerk der empirischen Untersuchung liegt jedoch nicht nur darauf, welche Methoden angewendet werden, sondern, wie die Methoden angewendet werden. Daher wurde insbesondere auch auf die **Intensität der Optimierungsmaßnahmen** und den **inhaltlichen Wert der Texte** geachtet (siehe Abbildung 16). Auffallend hierbei ist, dass die Webseiten, die sich objektiv mit dem Suchmaschinen-Optimierungswettbewerb beschäftigen, keine offensichtlichen Optimierungsmaßnahmen einsetzen, und dennoch erfolgreich in den drei betrachteten Suchmaschinen gelistet sind. Zu diesen Seiten zählt der Wikipedia-Artikel, die Heise News Seite und das Blog von COM.POSITUM.

Der Großteil der übrigen Webseiten hat mehr oder weniger einfallsreiche Geschichten über das Fabelwesen namens Hommingberger Gepardenforelle zum Thema. Für einen menschlichen Besucher haben diese Inhalte, wenn überhaupt, nur Unterhaltungswert. Die Ausnahme bildet hierbei die Webseite der Universität Kassel: Hier wurde die Suchphrase mit dem Ziel, dass die Suchbegriffe im Text in einer ausreichend hohen Dichte vorkommen, ohne Rücksicht auf den inhaltlichen Zusammenhang in einen ausgewählten Text eingefügt. Es entstand ein sehr langer Text ohne Sinn und somit auch ohne Nutzen für einen Leser.

	On-The-Page Kriterien	Off-The-Page Kriterien	Intensität der Optimierung	inhaltliche Beschreibung
ikipedia.org/ ngberger_ elle)	Keywordplatzierung: <title>-Tag, Meta-Tag keywords, <h1>, und <i> Schlagwortdichte: 0,90 %	Verlinkung: 90 Links, davon 4 externe Backlinks: 51 (G), 543 (Y), 1250 (M) Besonderheit: Inhaltsverzeichnis mit Textanker	keine offensichtliche Optimierung	Berichterstattung über den Wettbewerb
mingberger- elle.zielbewus	Keywordplatzierung: <title>-Tag, Meta-Tags keywords und description, <h1>, , alt-Texte bei Bildern, title-Attribute bei Links Schlagwortdichte: 2,50 %	Verlinkung: 23 Links, 6 externe Backlinks: 124 (G), 1360 (Y), 1430 (M)	Optimierung erkennbar, jedoch nicht sehr intensiv	Erfundene Texte zum Phantasiewesen Hommingberger Gepardenforelle
.danielgal.de/ rger- elle/ rger- elle.html)	Keywordplatzierung: <title>-Tag, Meta-Tags keywords und description, <h1>, <h2>, , alt- Texte und title-Attribute bei Bildern Schlagwortdichte: 3,90 %	Verlinkung: 18 Links, keiner davon extern Backlinks: 26 (G), 343 (Y), 937 (M)	stark optimierter Text und unzulässige Methoden (vesteckter Layer)	Erfundene Texte zum Phantasiewesen Hommingberger Gepardenforelle
usch mingberger- elle.malkusch.	Keywordplatzierung: <title>-Tag, Meta-Tag keywords, <h1>, <h2>, <h3>, , , <bd>, <abbr>, <var> Schlagwortdichte: 3,70 %	Verlinkung: 12 Links, keiner davon extern Backlinks: 170 (G), 1080 (Y), 1335 (M) Besonderheit: verlinkte Dateien in Ordern mit aussagekräftigem Namen	stark optimierter, langer Text, Fokus auf Textauszeichnung	Erfundene Texte zum Phantasiewesen Hommingberger Gepardenforelle
ssseite mingberger- elle.de)	Keywordplatzierung: <title>-Tag, Meta-Tags keywords und description, <h1>, <h3>, , <i>, alt-Text von Bildern, Kommentar Schlagwortdichte: 4,40 %	Verlinkung: 9 Links, keiner davon extern Backlinks: 114 (G), 5720 (Y), 36082 (M) Besonderheit: Kopie der Startseite (index_2.html), seit 15. April nicht aktualisiert	Optimierung erkennbar, jedoch nicht sehr intensiv	Erfundene Texte zum Phantasiewesen Hommingberger Gepardenforelle
assel .kde.cs.uni- re/ss2005/g	Keywordplatzierung: <title>-Tag, Meta-Tag keywords, <h1> Schlagwortdichte: 2,60 %	Verlinkung: 43 Links, 4 externe Backlinks: 30 (G), 1250 (Y), 142 (M)	stark optimierter, langer Text, Fokus auf Keyworddichte und eingehende Links	Zum überwiegenden Teil sinnloser Text

Tabelle 3: Zusammenfassung der Ergebnisse der qualitativen Untersuchung.

Webseite	On-The-Page Kriterien	Off-The-Page Kriterien	Intensität der Optimierung	inhaltliche Beschreibung
Heise News (http://www.heise.de/newsticker/meldung/58647)	Keywordplatzierung: <title>-Tag, Meta-Tag keywords, <h2> Schlagwortdichte: 0,50 %	Verlinkung: 103 Links, 9 externe Backlinks: 89 (G), 824 (Y), 1496 (M)	keine offensichtliche Optimierung	Berichterstattung über den Wettbewerb
Sitepoint (http://www.hommingberger-gepardenforelle.net)	Keywordplatzierung: <title>-Tag, Meta-Tags title, keywords, description, abstract, <h1>, <h2>, , alt-Text (Bilder), title-Attribut (Links) Schlagwortdichte: 2,50 %	Verlinkung: 27 Links, 6 externe Backlinks: 150 (G), 3030 (Y), 2370 (M) Besonderheit: Keywords in Dateinamen der Bilder	Optimierung deutlich erkennbar; Fokus auf Meta-Angaben und viel Text	Erfundene Texte zum Phantasiewesen Hommingberger Gepardenforelle
Tobias Mast (http://www.hommingberger-gepardenforelle-page.de)	Keywordplatzierung: <title>-Tag, Meta-Tags title, description, abstract, keywords, <h1>, und title-Attribut bei einem Bild Schlagwortdichte: 3,30 %	Verlinkung: 20 Links, 4 externe Backlinks: 537 (G), 5820 (Y), 1490 (M) Besonderheit: Kopie der Startseite (index_2.html), Keywords in Dateinamen	stark optimierter, vergleichsweise kurzer Text	Erfundene Texte zum Phantasiewesen Hommingberger Gepardenforelle
Teresia Kuhr (http://gepardenforelle.terrikay.de)	Keywordplatzierung: <title>-Tag, Meta-Tags keywords, description, author, copyright, page-type, page-topic, <h1>, <h2>, <h3>, , alt-Texte und title-Attribute bei Bildern Schlagwortdichte: 4,80 % Besonderheit: Codierungsfehler und fehlende head- und body-Tags	Verlinkung: 19 Links, 1 externe Backlinks: 49 (G), 3700 (Y), 235 (M) Besonderheit: Kopie der Startseite (index_2.html)	stark optimierter, vergleichsweise kurzer Text	Erfundene Texte zum Phantasiewesen Hommingberger Gepardenforelle
COM.POSITUM (http://www.hommingberger-gepardenforelle-blog.de)	Keywordplatzierung: <title>-Tag, Meta-Tags keywords, description, <h1>, <h3> , title-Attribute bei Links Schlagwortdichte: 1,80 %	Verlinkung: 97 Links, 19 externe Backlinks: 408 (G), 6150 (Y), 3951 (M) Besonderheit: Dateinamen mit Suchbegriffen	keine offensichtliche Optimierung	Texte über den Wettbewerb und Beiträge über Suchmaschinen-Optimierung

Unter den Seiten, die frei erfundene Texte präsentieren, scheinen die Methoden der Suchmaschinen-Optimierung von den teilnehmenden Privatpersonen, wie Daniel Gal, Markus Malkusch, Teresia Kuhr und Tobias Mast, intensiver angewendet worden zu sein. Einerseits äußert sich das durch die höheren Schlagwortdichten, andererseits erkennt man insbesondere an den On-The-Page Maßnahmen, dass sie eher bereit waren, die Qualitätsrichtlinien der Suchmaschinen zu verletzen. Beispiele hierfür finden sich insbesondere auf den Seiten von Daniel Gal (siehe Abbildung 11) und von Tobias Mast (siehe Abbildung 15). Die Webseiten, die von Agenturen erstellt worden sind, aber auch die offizielle Wettbewerbsseite setzen im Vergleich dazu die zur Verfügung stehenden Mittel dezent, also weniger intensiv, ein.

4.5 Schlussfolgerungen

Aus den Ergebnissen der Analyse kann man nur schwerlich auf all jene Faktoren schließen, die eine Webseite aufweisen muss, um in den Ergebnisseiten der Länderinterfaces von Google, Yahoo! und MSN unter die ersten 20 Ergebnisse zu gelangen. Dafür sind die Suchmaschinenalgorithmen zu komplex und vor allem der überwiegende Teil der Off-The-Page Kriterien zu wenig nachvollziehbar, sodass selbst automatisierte quantitative Untersuchungen, wie sie etwa der Heise Verlag durchgeführt hat, keine verlässlichen Ergebnisse liefern.

Vielmehr kann man aus der vorliegenden Untersuchung, in der elf ausgewählte Webseiten qualitativ analysiert wurden, ableiten, welche Rolle Suchmaschinen-Optimierung für den Erfolg einer Webseite spielt. Auf Basis der vorliegenden Informationen gilt es herauszufinden, ob es nur die Methoden der On-The-Page Optimierung und deren Anwendung sind, die für den Erfolg einer Webseite in den Ergebnislisten der Suchmaschinen verantwortlich sind, oder ob auch der Inhalt in irgendeiner Form auf die Platzierung in den Suchmaschinen Einfluss nimmt. Damit leistet die vorliegende Arbeit einen Beitrag für das Verstehen von Suchmaschinen-Optimierung, der über die Kenntnis der Methoden und die Vorlieben der Suchmaschinenalgorithmen hinausgeht: Die Bedeutung des Inhalts im Verhältnis zu den Methoden der Suchmaschinen-Optimierung steht im Mittelpunkt der folgenden Ausführungen.

Die Ergebnisse zeigen, dass sich die untersuchten Webseiten aufgrund mehrerer gemeinsamer Eigenschaften in vier – von theoretisch neun möglichen –

Typen gruppieren lassen (siehe auch Abbildung 16). Der erste Typ ist jener der **Seiten mit ansprechenden Inhalten, jedoch ohne offensichtlicher On-The-Page-Optimierung** und umfasst den Beitrag von Wikipedia, der Seite von Heise News und das Weblog der Agentur COM.POSITUM. In Betracht der großen Konkurrenz an stark optimierten Webseiten, die in einem SEO-Wettbewerb herrscht, hat man nicht erwarten können, dass es überhaupt Seiten geben wird, die es ohne On-The-Page-Optimierung an die Spitze der Ergebnisseiten schaffen. Für diesen überraschenden Erfolg müssen Off-The-Page Faktoren, über die im Rahmen der Analyse zu wenige verlässliche Daten für eine verlässliche Analyse zur Verfügung standen, verantwortlich sein. Hierzu zählen Kriterien, wie Link Popularität³⁹ und Klick Popularität, da die Seiten in Bezug auf den Wettbewerb interessante Inhalte bieten, von denen es denkbar ist, dass sie häufig verlinkt (vgl. Link Popularität) und häufig besucht (vgl. Klick Popularität) worden sind.

Den zweite Typus von Seiten stellen die **Seiten mit Phantasiertexten und vergleichsweise schwacher Optimierung** dar. Neben der offiziellen Wettbewerbsseite zählen hierzu die Seiten von Sitepoint und die Zielbewusst-Seite. Nicht nur aufgrund der relativ geringen Höhe der Schlagwortdichte, scheinen auch bei dieser Gruppe nicht die On-The-Page Faktoren für den Erfolg entscheidend gewesen zu sein. Wiederum ist anzunehmen, dass die Off-The-Page Faktoren für die guten Platzierungen in den Suchmaschinen verantwortlich sind. Insbesondere bei der Wettbewerbsseite ist anzunehmen, dass viele Betreiber von teilnehmenden Seiten als Referenz an den Wettbewerb einen Backlink gesetzt haben, was durch die Angaben, die MSN liefert, bestätigt wird. Auch in Bezug auf die Sitepoint- und die Zielbewusst-Seiten könnte eine hohe Anzahl der Backlinks aus dem Linknetzwerk, das die Agentur betreibt, die geringe Intensität der On-The-Page Optimierung kompensiert haben. Beweise für diese Annahme gibt es jedoch nicht. Der Inhalt kann jedenfalls kaum ausschlaggebend für das Setzen eines Links auf diesen Seiten sein, da es sich durchwegs um erfundene Geschichten handelt, die sich kaum von anderen Wettbewerbsbeiträgen unterscheiden.

Gliedert man die Seiten nach dem Ersteller, was kurz angedacht wurde, sind im Rahmen der Untersuchung **stark optimierte Seiten mit erfundenen Texten** besonders für Privatpersonen charakteristisch. Hier werden die Mög-

³⁹ Besonders für das Weblog von COM.POSITUM, scheint die hohe Anzahl an Backlinks darauf hinzudeuten.

lichkeiten der On-The-Page Optimierung teilweise bis an die Grenzen, die die Suchmaschinen mit ihren Qualitätsrichtlinien setzen, ausgeschöpft. Dies äußert sich durch besonders hohe Keyworddichten, aber auch durch die Platzierung der Suchbegriffe in besonders vielen Textauszeichnungen (siehe Kapitel 4.4). Auf die Verlinkung wird offensichtlich weniger Wert gelegt, auch wenn die Seiten von Tobias Mast in Google besonders viele Backlinks zu haben scheint.

Abschließend bleibt noch der Wettbewerbsbeitrag der Universität Kassel. Auf der indexierten Seite wird nur **wenig sinnvoller und schwach optimierter Text** präsentiert. Der Grund für die guten Platzierungen scheint zudem auch nicht in der Verlinkung zu liegen, auch wenn der Betreiber der Webseite dazu aufgerufen hat, möglichst viele Links auf die Seite zu setzen. Es könnte vielmehr in der Länge des Textes begründet sein, worin die Suchmaschinen einen Indikator für Qualität sehen.

Texte ohne Sinn		Universität Kassel	
Phantasietexte		Wettbewerbsseite Zielbewusst Sitepoint	Markus Malkusch Daniel Gal Teresia Kuhr Tobias Mast
objektive Berichterstattung	COM.POSITUM Wikipedia Heise News		
	keine Optimierung	schwache Optimierung	starke Optimierung

Abbildung 16: Grafische Darstellung der Webseiten in Bezug auf inhaltliche Beschreibung und Intensität der Suchmaschinen-Optimierung.

Für Überprüfung der formulierten Hypothese „Webseiten, die in den Ergebnislisten von Suchmaschinen besonders gut platziert sind, zeichnen sich sowohl durch ausgeprägte Anwendung von On-The-Page Methoden, als auch durch inhaltliche Qualität aus.“ bedeutet das folglich, dass die Hypothese nicht verifiziert werden kann. Die Ergebnisse der Untersuchung deuten nämlich darauf hin, dass einerseits Webseiten, die keine erkennbaren Anstrengungen in Maßnahmen bezüglich On-The-Page Optimierung investiert haben und sich ausschließlich auf den Inhalt konzentriert haben, von Suchmaschinen gut bewertet werden. Andererseits findet sich auch ein Beispiel dafür, dass eine Seite ohne sinnvollen Text aufgrund von gezielten Suchmaschinen-Optimierungsmaßnahmen besonders erfolgreich ist.

Die elf untersuchten Webseiten unterscheiden sich also in Bezug auf ihren Inhalt der Intensität der Anwendung von Methoden der On-The-Page Optimierung teilweise beträchtlich voneinander, während sich die Methoden selbst, die in Kapitel 4.4 zusammengefasst dargestellt worden sind, häufig ähneln. Das Raster in Abbildung 16 zeigt zudem, dass die Seiten in Bezug auf die Intensität der Optimierung annähernd gleich verteilt sind: Es finden sich drei Seiten ohne erkennbare Optimierungsmaßnahmen und je vier schwach, beziehungsweise stark optimierte Seiten. Inhaltlich haben mehr als die Hälfte der Seiten, nämlich sieben, Phantasietexte über die Hommingberger Gepardenforelle präsentiert. Im Gegensatz dazu waren nur drei Seiten mit objektiver Berichterstattung und eine Seite mit einem Text ohne Sinnzusammenhang zu finden.

Geht man davon aus, dass die Seiten von Wikipedia und die Heise News Seite nicht mit der Absicht erstellt worden sind, am Wettbewerb teilzunehmen, erkennt man, dass hauptsächlich optimierte Seiten ohne sinnvollen Inhalt und besonderen Nutzen für den Leser im Rahmen des Suchmaschinen-Optimierungswettbewerbes erfolgreich waren. Aus den Beispielen, die der Suchmaschinen-Optimierungswettbewerb im Rahmen dieser qualitativen Untersuchung hervorgebracht hat, kann man also schließen, dass für den kurzfristigen Erfolg einer Webseite, wie er insbesondere bei einem SEO-Wettbewerb angestrebt wird, eine intensive Anwendung von On-The-Page Methoden bis in die Grauzonen des Erlaubten durchaus Erfolg verspricht. Auf lange Sicht, insbesondere außerhalb von SEO-Wettbewerben, scheint jedoch die Gefahr groß zu sein, bei einer Änderung der Suchmaschinen-Algorithmen, als Spam zu gelten und gänzlich aus den Suchmaschinenresultaten zu verschwinden. Einzig-

artige und abwechslungsreiche Inhalte hingegen, zum Beispiel im Beitrag von Wikipedia und Heise News oder im Weblog der Seiten von COM.POSITUM, ziehen viele Besucher und viele Links an, was sich positiv auf die Link Popularität und Klick Popularität auswirkt und die Platzierung in den Suchmaschinen positiv beeinflusst.

Besonders vor dem Hintergrund, dass die On-The-Page Kriterien vermehrt, im Sinne von anfrageabhängigen Faktoren (siehe Seite 8), dazu dienen, die Relevanz eines Suchbegriffes für die gestellte Suchabfrage festzustellen, während anhand der Off-The-Page Faktoren unabhängig von der Anfrage die Reihung der Ergebnisse nach Qualität vorgenommen wird, kommt der Qualität der Inhalte wieder größere Bedeutung für die Platzierung in den Suchmaschinen zu.

Um einen langfristigen Erfolg in den Ergebnislisten der Suchmaschinen sicherzustellen, spielen also einerseits die Methoden der Suchmaschinen-Optimierung eine Rolle, andererseits jedoch auch der Inhalt der Seite. Hierbei eignet sich die gezielte Anwendung von On-The-Page Methoden dafür, jene ausgewählten Suchbegriffe hervorzuheben, für die die Suchmaschine die optimierte Webseite als relevant bewerten soll. Man nimmt also darauf Einfluss, zu welchen Suchbegriffen man in der Suchmaschine gefunden wird.

Auf die Platzierung in den Ergebnislisten der Suchmaschinen kann man jedoch keinen unmittelbaren Einfluss nehmen, da für die Qualitätsbewertung jene Off-The-Page Faktoren herangezogen werden, die außerhalb der Kontrolle des Webmasters liegen, wie Link Popularität oder Klick Popularität. Sofern man also kein gut funktionierendes Linknetzwerk betreibt, kann man nur durch interessante Inhalte, die zahlreiche Besucher und Backlinks anlockt, zumindest indirekt die Platzierung auf den Ergebnisseiten beeinflussen.

Die Entwicklung bei den Suchmaschinenbetreibern geht jedoch rasch voran. Insbesondere, um Verstöße gegen die Qualitätsrichtlinien besser zu erkennen, werden häufig die Suchalgorithmen verändert. Diese Veränderungen haben oft auch für Seiten, die sich keiner unzulässigen Methoden bedienen, Auswirkungen. Auch ist in Zukunft mit neuen Konzepten zu rechnen: Während Google offenbar weiterhin in technische Infrastruktur investiert, um sein Netzwerk weiter zu vergrößern, setzt Yahoo!, wie die kürzlich getätigten Firmenakquisitionen vermuten lassen, auf soziale Netzwerke im Internet, etwa von Portalen, Weblogs, Wikis und ähnlichem (vergleiche Web 2.0). Die große Unbekannte ist jene europäische Suchmaschine, die von manchen französischen und deutschen

Politikern gefordert wird, um die Vorherrschaft und Abhängigkeit von amerikanischen Suchmaschinen zu brechen. Sollte sie realisiert werden, könnte das einerseits eine Veränderung am Suchmaschinenmarkt nach sich ziehen, andererseits sind auch von dieser Suchmaschine innovative Konzepte zu erwarten.

Wie lange die Methoden, die im Rahmen dieser Arbeit vorgestellt wurden, Gültigkeit haben werden ist zu diesem Zeitpunkt noch nicht absehbar. Denn, wie das Internet im Gesamten, ist Suchmaschinen-Marketing ständigen Veränderungen unterworfen. Für den Betreiber einer Internetseite wird es wichtig sein, die Entwicklungen am Suchmaschinenmarkt und die Innovationen der Suchmaschinen zu beobachten und kritisch auf ihre Tauglichkeit zu prüfen. Für alle Beteiligten, seien es die Betreiber der Suchmaschinen, die Ersteller von Webseiten oder die Benutzer, die mit Suchmaschinen nach Informationen suchen, ist es jedenfalls wünschenswert, dass die Qualität Inhalts für den Erfolg einer Webseite in Zukunft eine noch gewichtigere Rolle spielt.

Literatur

- Aaberg, Ingvar (2006):** Search Engine Technologies. Status and Directions. $\langle$ URL: <http://www.muenchner-kreis.de/presentations/ondemand.php?path=presentations/video&media=1&size=256&ppt=RMPM00309> $\rangle$ – Zugriff am 9. Mai 2006.
- Bager, Jo (2005a):** Turbolift ins Ungewisse. Manipulationsmaschinen bei Suchmaschinen. in: c't - Magazin für Computertechnik, Nr. 9, 164 ff..
- Bager, Jo (2005b):** Virtuelle Schwärme. Erste Ergebnisse des c't-Suchmaschinen-Optimier-Wettbewerbs. in: c't - Magazin für Computertechnik, Nr. 10, 43.
- Bager, Jo (2006):** Dicke Fische. Zweite Runde des c't-Suchmaschinenoptimierungswettbewerbs. in: c't - Magazin für Computertechnik, Nr. 1, 76.
- Battelle, John (2005):** The Search: How Google and its Rivals rewrote the Rules of Business and transformed our Culture. London: Penguin Books.
- Bruce Clay Inc. (2006):** Search Engine Chart. $\langle$ URL: <http://www.bruceclay.com/searchenginechart.pdf> $\rangle$ – Zugriff am 9. Mai 2006.
- Bundesverband Digitale Wirtschaft (2004):** PG Suchmaschinen-Marketing des AK Performance Marketing. $\langle$ URL: http://www.bvdw.org/ww/de/7_pub/themen_neu/performance_marketing/suchmaschinenmarketing.cfm $\rangle$ – Zugriff am 9. Mai 2006.
- Garfinkel, Simson (2004):** Geheimniskrämerei bei Google? $\langle$ URL: <http://www.heise.de/tr/artikel/print/46973> $\rangle$ – Zugriff am 9. Mai 2006.
- Glöggler, Michael (2003):** Suchmaschinen im Internet. Berlin: Springer.
- Google (2005a):** Richtlinien für Webmaster. $\langle$ URL: <http://www.google.at/intl/de/webmasters/guidelines.html> $\rangle$ – Zugriff am 9. Mai 2006.
- Google (2005b):** Mich verwundert die Rangfolge meiner Website. $\langle$ URL: <http://www.google.at/intl/de/webmasters/4.html> $\rangle$ – Zugriff am 9. Mai 2006.
- Hannemann, Ulf/Jutzi, Sebastian/Körner, Andreas (2005):** Das große Googeln. in: Focus, Nr. 46, 112 – 116.

- Heise Online (2006):** c't-SEO-Wettbewerb-Ergebnisse. $\langle$ URL: <http://www.heise.de/ct/SEO-Wettbewerb/ergebnisse2.shtml> $\rangle$ – Zugriff am 9. Mai 2006.
- Karzauninkat, Stefan (2002):** Die Suchfibel. 3. Auflage. Leipzig: Klett Verlag $\langle$ URL: <http://www.suchfibel.de/> $\rangle$ – Zugriff am 9. Mai 2006.
- Kotler, Philip/Bliemel, Friedhelm (2001):** Marketing-Management: Analyse, Planung, Umsetzung und Steuerung. 10. Auflage. Stuttgart: Schäffer-Poeschel.
- Krug, Steve (2002):** Don't make me think!. Web Usability - Das intuitive Web. Bonn: mitp-Verlag.
- Lewandowski, Dirk (2005):** Web Information Retrieval. Technologien zur Informationssuche im Internet. Frankfurt am Main: DGI-Schrift (Informationswissenschaft 7).
- Meier, Berthold (2006):** Hinweise und Links für InformationsspezialistInnen. $\langle$ URL: <http://people.freenet.de/biblio/IuD/intsmueb.htm> $\rangle$ – Zugriff am 9. Mai 2006.
- Michael, Alex/Salter, Ben (2003):** Marketing through Search Optimization. How to be found on the Web. Burlington: Butterworth-Heinemann (Elsevier).
- MSN (2006):** Richtlinien für eine erfolgreiche Indizierung. $\langle$ URL: http://search.msn.at/docs/siteowner.aspx?t=SEARCH_WEBMASTER_REF_GuidelinesforOptimizingSite.htm $\rangle$ – Zugriff am 9. Mai 2006.
- Münz, Stefan (2005):** SELFHTML Version 8.1.1. $\langle$ URL: <http://de.selfhtml.org> $\rangle$ – Zugriff am 9. Mai 2006.
- Ott, Stephan (2006):** Links And Law - Leitfaden zur Manipulation von Suchergebnissen. $\langle$ URL: <http://www.linksandlaw.de/suchmaschinen-leitfaden-zur-Manipulation-von-Suchergebnissen.htm> $\rangle$ – Zugriff am 9. Mai 2006.
- Seda, Catherine (2004):** Search Engine Advertising. Buying your Way to the Top to increase Sales. Indianapolis: New Riders Publishing.
- Speck, Prof. Dr. Hendrik (2006):** Socio-Political Observations. $\langle$ URL: <http://www.muenchner-kreis.de/presentations/ondemand.php?>

path=presentations/video&media=1&size=256&ppt=RMPM00326⟩ –
Zugriff am 9. Mai 2006.

Thurrow, Shari (2003): Search Engine Visibility. Indianapolis: New Riders Publishing.

Yahoo! (2005): Wie lauten Ihre Leitlinien zu Spam? ⟨URL:
[http://eur.help.yahoo.com/help/de/ysearch/deletions/
deletions-05.html](http://eur.help.yahoo.com/help/de/ysearch/deletions/deletions-05.html)⟩ – Zugriff am 9. Mai 2006.

Lebenslauf des Autors

STECKBRIEF

Name	Hans-Peter Bauer
Staatsbürgerschaft	Österreich
Geburtstag	31. August 1983 in Wiener Neustadt
Anschrift	Walbersdorf, Bahnstraße 8, 7210 Mattersburg
eMail	hpb@gmx.at
Website	http://hpb.at.tt

AUSBILDUNG

- 1989 – 1993** Volksschule Mattersburg
- 1993 – 2001** Klemens Maria Hofbauer-Gymnasium Katzelsdorf
- 2002 – 2006** Fachhochschul-Studiengang Informationsberufe Eisenstadt

PRAKTIKA (AUSWAHL)

- **NÖN/BVZ Mattersburg (Juni 2003 und August 2003)**
Lokaljournalismus in der Region Wulkatal, Bezirk Mattersburg.
- **Wirtschaftskammer Burgenland, Eisenstadt (September 2003 und August 2004)**
Assistenz des Spartengeschäftsführers für Gewerbe und Handwerk.
- **eTel, Eisenstadt (Juli 2004)**
PHP Programmierung in Verbindung mit MySQL.
- **T-Mobile Austria, Wien (Juli und August 2005)**
PHP Programmierung in Verbindung mit Oracle 9 Datenbanken.
- **LKW Walter Internationale Transportorganisation AG (September 2005 bis Jänner 2006)**
Mitarbeit im Konzernmarketing.

AUSBILDUNGSPROJEKTE

- **Einführung eines Dokumentenmanagement Systems (SS 2003)**
- **Konzipierung und Realisierung eines Webradios (WS 2003/2004 und SS 2004)**
- **Webauftritt für den Fachhochschul-Studiengang Informationsberufe (WS 2004/2005 und SS 2005)**