



Jo Bager, Jan Becker-Fochler

Dicke Fische

Zweite Runde des c't-Suchmaschinenoptimierungswettbewerbs

Der c't-SEO-Wettbewerb erlaubt Einblicke in die Praxis der Suchmaschinenoptimierung, die sowohl für Benutzer der Suchdienste als auch für Webmaster interessant sind. Viele Webmaster haben bis zum zweiten Stichtag um die guten Platzierungen ihrer Sites gekämpft.

Bis zum 14. April dieses Jahres war die Gattung der Hommingberger Gepardenforelle gänzlich unbekannt. Einen Tag später erschienen die c't 9/05 und eine News-Meldung auf heise online, die zum c't-Suchmaschinenoptimierungswettbewerb aufrufen [1]. Aufgabe: Webseiten zu veröffentlichen, die in den Trefferlisten von Google.de, Yahoo.de, MSN.de und Seekport.de für die Abfrage *Hommingberger Gepardenforelle* Spitzenpositionen belegen. Nach einer Bestandsaufnahme am 15. Mai war der 15. Dezember der Stichtag, an dem wir prüften, wer sich über längere Zeit auf den vorderen Plätzen behaupten konnte.

Beim c't-Wettbewerb ging es nur um die Ehre. Dennoch fand der Wettstreit ein enormes Medienecho und regen Zulauf. Wie viele Sites sich beteiligt haben, lässt sich nicht sagen – es gab keine Anmeldeprozedur. Es müssen aber etliche Hundert gewesen sein: Die Site *vwp-online.de* [5], die für jeden Tag des Wettbewerbs die hundert bestplatzierten Sites festgehalten hat, zählte alleine 185 Domains mit dem Namensbestandteil *Gepardenforelle*.

Offenbar erhofften sich etliche Suchmaschinenoptimierer (search engine optimizer, SEOs) durch ein gutes Ergebnis im Wettbewerb auch Werbung für ihr Unternehmen. So investierten einige Webmaster bares Geld für eine gute Position. Mit Anzeigen in Google AdWords warben sie für Links auf ihre Homepage.

Mit harten Bandagen

Schon kurz nach Start des Wettbewerbs zeigte sich, dass mit harten Bandagen gearbeitet wurde. So schrieb Matthias Meier, Betreiber von *hommingberger-gepardenforelle.net*, eine geharnischte E-Mail an die Spiegel-Online-Redaktion. Diese hatte gewagt, einen Artikel über den Wettbewerb zu veröffentlichen und andere Wettbewerbsites, aber nicht seine, zu verlinken.

Dass vor diesem Hintergrund nicht immer mit sauberen Mitteln gekämpft wurde, liegt auf der Hand. Ein Website-Betreiber versuchte zum Beispiel, seine Seite kurz vor Ablauf der zweiten Deadline durch einen Kunstgriff auf die gute Position der *heise-Newsticker*-meldung (Auf der ersten Trefferseite bei Google und

Yahoo) zu mogeln. Der Trick nutzt einen Fehler von Google und Yahoo aus und ist in SEO-Kreisen auch als 302 Hijacking oder URL Hijacking bekannt.

Dabei richtet der Betreiber einer schlecht positionierten Seite auf seinem Webserver eine Weiterleitung (*http-Redirect*) auf die Adresse einer besser platzierten Seite ein. Mit dem Fehlercode 302 lügt der weiterleitende Webserver dabei dem Browser oder dem Suchmaschinen-Robot vor, dass die angefragte Datei vorübergehend an der besser positionierten Adresse zu finden wäre. Google und Yahoo zeigen daraufhin in einigen wenigen Fällen statt der URL der besser platzierten Seite diejenige der Weiterleitungsseite an. Kurz vor der zweiten Deadline hat der Trick bei Yahoo geklappt.

Linkzählung

Die Tabelle auf Seite 78 hilft, eine Hand voll SEO-relevanter Faktoren der jeweils ersten zehn Resultate aller Suchmaschinen zu vergleichen. Unsere eigene Site, *www.hommingberger-gepardenforelle.de*, läuft zwar im Wettbewerb außer Konkurrenz mit. Wir haben sie für Vergleichszwecke aber trotzdem in der Tabelle belassen.

Die Tabelle enthält als einziges Kriterium, das nur Eigenschaften der Seite selbst berücksichtigt (*On the Page*), die so genannte *Keyword-Dichte*, also das Verhältnis der Schlüsselwörter zum gesamten *Body-Text* (zur *On-the-Page-Optimierung* siehe auch [1]).

Daneben gewinnen externe, so genannte *Off-the-Page-Faktoren* immer mehr an Bedeutung. Wir haben mit Hilfe des Programms *Backlink Spider* die Verlinkung genauer unter die Lupe genommen (siehe *Soft-Link*). Es ermittelt nicht nur die Anzahl

der Links auf eine Seite (*Backlinks*), sondern zählt auch die *Domains* und die *IP-Adressen* durch, von denen verlinkt wird – weitere Qualitätsmerkmale.

Backlink Spider geht Seite für Seite die Trefferlisten von Google, Yahoo und MSN durch und protokolliert alle Treffer, die einen Link auf die gesuchte Seite gesetzt haben. Diese Methode hält sicherlich nicht wissenschaftlichen Ansprüchen stand, da Suchmaschinen nicht alle Seiten aus ihrem Index ausliefern, die der Suchanfrage entsprechen. Um den Stellenwert von Sites zu vergleichen, reichen die Ergebnisse aber allemal.

Als besonderes Schmäckerl schlüsselt der *Backlink Spider* auf, wie viele Links von Sites mit welchem *Page Rank* stammen, Googles Einschätzung der Wichtigkeit einer Seite. Dieser *Google-interne Wert* hat zwar eigentlich nur bei Google Gewicht – anderen Suchmaschinen steht er für das Ranking nicht zur Verfügung. Das Programm bezieht ihn über eine Programmierschnittstelle bei Google.

Da sich das Ranking aller Suchdienste in vielen Punkten angleicht, reicht er für eine grobe Klassifizierung aus: Sites mit einem hohen *Page Rank* dürften von allen Suchdiensten als wichtig, als so genannte *Authorities* angesehen werden. Der *Page Rank* einer einzelnen Site lässt sich mit der *Google Toolbar* für den *Internet Explorer* und *Firefox* ermitteln – siehe *Soft-Link*. Für diese optionale Funktion sendet die *Toolbar* die besuchten *URLs* an die *Google-Server*.

Außerdem listet die Tabelle einen Wert für den so genannten *Klick Rank*. SEOs haben herausgefunden, dass der *Traffic*, also der *Besucherstrom* zu einer Site, sich auch maßgeblich auf die Einstufung in den Trefferlisten auswirken kann. Mit der *Toolbar*

hat Google ein Mittel, um Besucherströme zu erfassen. Redirects, ein weiteres Hilfsmittel, das sich für die Traffic-Protokollierung eignet, findet man bei allen Suchmaschinen.

Da wir auf die Werte der Suchmaschinenbetreiber nicht zugreifen können, enthält die Tabelle die Werte der Firma Alexa. Sie wertet Klicks der Nutzer ihrer Toolbar aus. Der Wert 1928 bei der Heise-Newstickermeldung bedeutet, dass heise.de die am 1928sten häufig besuchte Site von Nutzern der Alexa-Toolbar ist.

Google

Google rules. Ein Großteil der Surfer benutzt den Primus für die Recherche, und einige Teilnehmer, etwa Matthias Meier, Betreiber von www.hommingberger-gepardenforelle.net, ziehen daraus die Konsequenz: „Ich optimiere nur für Google. Alles andere kann man vernachlässigen.“

Ein wichtiges Kriterium für das Ranking bei Google ist der Page Rank. Die Tabelle zeigt allerdings sehr schön, dass er nicht ausschließlich die Reihenfolge der Treffer bestimmt, sondern offenbar von anderen Faktoren „überstimmt“ werden kann: Der Wikipedia-Artikel und www.hommingberger-gepardenforelle-blog.de sind vor Seiten mit einem höheren Page Rank positioniert.

Bei beiden Sites dürfte ein On-the-Page-Faktor ausschlaggebend gewesen sein, der sich nicht gut in der Tabelle erfassen lässt: Sie sind auch in der Zeit zwischen den Stichtagen regelmäßig aktualisiert worden. Das Blog, aufgesetzt vom Auszubildenden Oliver Baumgart der Web-Agentur COM.POSITUM, war zwar nicht das einzige Weblog im Wettbewerb, hat aber am erfolgreichsten mit guten Einfällen immer wieder neue Inhalte generiert (und damit auch Links auf sich gezogen). So veranstaltete es einen Poetry-Slam-Wettbewerb, bei dem es darum ging, Gepardenforellen-Gedichte zu reimen. Mit solchen Aktionen hat sich das Blog von Position 75 im April auf den 9. Platz am 15. Mai und den 6. Platz am 15. Dezember hochgearbeitet.

Für regelmäßig aktualisierte Inhalte haben auch die Teilnehmer des Seminars „Internet-Suchmaschinen“ des Fachgebietes

Wissensverarbeitung an der Universität Kassel auf ihrer Seminar-Homepage gesorgt. Da es in dem Seminar ohnehin um die Analyse von Spam-Methoden gehen sollte, wurde die Wettbewerbsteilnahme kurzerhand zur praktischen Übung.

Die Seminar-Website zeigt aber auch anschaulich, dass von Google geschätzte, aktuelle Inhalte für den Surfer ohne jeden Wert sein können. Die Studenten und wissenschaftlichen Mitarbeiter haben das Programm SCigen – An Automatic CS Paper Generator angeworfen [7] und sich automatisch ein pseudowissenschaftliches Nonsense-Paper generieren lassen. Das englischsprachige Dokument haben sie durch Babelfish ins Deutsche übersetzen lassen und anschließend so viele Substantive ersetzt, bis das Suchwort häufig genug vorkam. Ehrlicherweise enthält die URL ihrer Seite den Begriff googlespam.

Haupterfolgsrezept der Seminarteilnehmer – neben der von Google vielen Uni-Websites zugebilligtem inhaltlich hohen Wertschätzung – ist die massive Verlinkung der Seite. Die Studenten haben ihre Kontakte spielen lassen und Webmaster anderer Institute um Linksetzungen gebeten. Auch die Uni-Pressestelle hat mitgespielt, eine Pressemitteilung veröffentlicht und für einen Link von der Universitäts-Homepage gesorgt. Selbst in der Lokalpresse wurde um Link-Spenden gebeten. Der Ertrag: Immerhin drei Backlinks von Sites mit Page

Anzeigen

[10€/Link & Ipad gewinnen?](#)
Pro **Hommingberger Gepardenforelle** die jetzt zu mir schwimmt!
www.ranking-check.de/Gepardenforell

[Gepardenforelle zum Wein](#)
SEOs gesucht
Hohe 21% Provision für Sie
www.ebrosia.com

[AdWords für Forellen?](#)
Wieviel Aufwand betreiben Sie für die **Hommingberger Gepardenforelle?**
hommingberger-geparden-forellen.de

[Die echte Gepardenforelle](#)
Hommingberger Gepardenforelle
Für Webmaster aller Art - Lecker!
www.wcheck.com/HommingbergerForelle

[Angeln in Hommingberg?](#)
Kinderferienlager & Jugendcamps in Deutschland und Europa
www.hommingberg.ferienlager.org

Einige Site-Betreiber haben sogar Geld für Links ausgegeben.

Rank 6 und 18 5er-Backlinks und ein achter Platz bei Google.

Dass es nicht nur wichtig ist, von irgendwelchen Sites mit hohem Page Rank verlinkt zu werden, sondern von den richtigen, lässt sich auch am Abschneiden der Site www.hommingberger-gepardenforelle.net zeigen. Am ersten Stichtag erstplatziert, sackte sie zwischenzeitlich unter den 60. Rang. Besitzer Matthias Meier führt das darauf zurück, dass er zwischenzeitlich zu viele Links von englischsprachigen Sites erhalten hat. Google hat seine Site daraufhin als englisch eingestuft – derzeit ein Kriterium für die Abwertung in den Trefferlisten der deutschen Google-Ausgabe.

Yahoo

Unterschieden sich die Ranking-Kriterien von Yahoo am ersten Stichtag noch wesentlich von denen Googles, so haben sich beide Suchmaschinen angeglichen. Yahoo legt mittlerweile offenbar auch wesentlich mehr Wert auf die Verlinkung. Bei der On-the-Page-Optimierung scheint Yahoo sogar strengere Maßstäbe anzulegen als Google. Keine Website in den Top Ten bei Yahoo hat eine Schlagwortdichte von mehr als 4,8 Prozent.

MSN

Aus Benutzersicht kann man die Ergebnisse von MSN nur als Katastrophe bezeichnen. So finden sich gleich drei Treffer einer Link-Farm unter den Top Ten. Wenn sich Sites mit Domains der Form www.hommingberger1gepardenforelle.de bis www.hommingberger4gepardenforelle.de fast ausschließlich untereinander verlinken, dann sollte das eine Suchmaschine heutzutage als Spam erkennen.

Als weitere Spam-Methode benutzen die Sites so genannten Webalizer Spam. Dabei sucht der Webmaster nach gut positionierten Webseiten, die eine Statistik der Webserver-Log-Software Webalizer veröffentlichen, die von MSN indexiert ist. Diese Domains lässt der Angreifer dann immer wieder per Skript aufrufen. Dem Server des „Opfers“ wird dabei immer die eigene Homepage als Referrer übergeben. Die Webalizer-Statistik der jeweiligen Domain generiert eine Liste der TOP 30 Referrer und verlinkt

diese. Spidert ein MSN-Bot diese Seiten, so wird der Link als Empfehlung für den Angreifer gezählt. Bei den anderen Suchdiensten schafft man es mit solchen Tricks nicht auf die vorderen Ränge.

Auch ist völlig schleierhaft, warum sich eine Site unter den Top Ten befindet, die von gerade einmal 25 anderen Websites verlinkt wird. Die Tatsache, dass die erste Ergebnisseite nur neun statt der eingestellten zehn Treffer enthielt, bestätigt den schlechten Eindruck. Alles in allem scheint MSN derzeit eher ein Spielfeld zu sein, um veraltete Spamming-Techniken auszuprobieren, denn ein ernst zu nehmender Suchhelfer.

Seekport

Die Ergebnisse von Seekport nehmen eine Sonderstellung ein, denn der Suchmaschinenbetreiber ist der einzige, der sich offen zur manuellen Filterung bekennt und sie auch konsequent anwendet – Seiten, die das Filter-Team als Spam einschätzt, fliegen aus dem Index.

Aus diesem Grund lässt sich der Einfluss von formalen Aspekten wie Backlinks oder der Schlagwortdichte schlecht einschätzen. Man kann nicht nachvollziehen, an welcher Stelle der Suchergebnisse der händische Eingriff endet und wo die vom Ranking-Algorithmus eingestufteten Sites beginnen. Für den Wettbewerb lässt sich beobachten, dass sich unter den Top Ten nur Seiten fanden, die dem Leser sinnvolle Inhalte präsentierten (sofern das bei einem virtuellen Wesen wie der Hommingberger Gepardenforelle überhaupt möglich ist).

Fazit

Der c't-Wettbewerb zeigt, dass Benutzer von Suchmaschinen die Ergebnisse von Google und Co. stets mit einer gewissen Portion Skepsis betrachten sollten. Auf den ersten Rängen der Trefferlisten tummeln sich häufig nicht die aus Nutzersicht besten und informativsten Sites, sondern diejenigen, deren Webmaster den meisten Aufwand für die Optimierung betrieben haben.

Webmaster können durch die Analyse der bestplatzierten Sites Rückschlüsse darauf ziehen, was sie tun müssen, um ihre Sites besser zu positionieren. Die Tabelle

c't-SEO-Wettbewerb, zweiter Stichtag, 15.12.

Rang/ Such- maschine	Suchmaschine/ URL	Schlagwort- Dichte	Google- PageRank	Anzahl Backlinks Domains	Anzahl Backlink-	Anzahl Backlink-Domains mit PR 7/6/5/4/3/2/1/0	Anzahl Backlink- IP-Adresen	Klick- Popularität
Google								
1	www.hommingberger-gepardenforelle.de	4,4 %	6	882	613	0/6/11/25/62/56/39/414	498	794 741
2	www.hommingberger-gepardenforelle.net	2,5 %	6	860	530	0/3/19/34/32/57/40/345	441	2 088 380
3	http://de.wikipedia.org/wiki/Hommingberger_Gepardenforelle	0,9 %	5	248	233	0/0/6/14/17/15/4/177	151	35
4	www.hommingberger-gepardenforelle-page.de	3,3 %	5	409	249	0/0/15/34/45/29/31/95	210	1 106 352
5	www.heise.de/newsticker/meldung/58647	0,5 %	6	752	539	0/0/10/35/44/52/29/369	432	1 928
6	www.hommingberger-gepardenforelle-blog.de	1,8 %	5	287	152	0/0/12/20/20/14/9/77	124	4 811 469
7	www.informationsarchiv.com/webkatalog/Hommingberger-Gepardenforelle/60967.shtml	5,6 %	6	123	96	1/1/15/19/14/16/5/25	88	145 670
8	www.kde.cs.uni-kassel.de/lehre/ss2005/googlespam	2,6 %	5	218	139	0/3/18/14/16/10/4/74	103	93 339
9	http://hommingberger-gepardenforelle.zielbewusst.de	2,5 %	5	487	288	0/2/19/27/24/48/36/132	236	2 346 495
10	www.danielgal.de/hommingberger-gepardenforelle/hommingberger-gepardenforelle.html	3,9 %	5	209	147	0/0/16/24/20/19/11/57	126	-
Yahoo								
1	http://hommingberger-gepardenforelle.zielbewusst.de	2,5 %	5	487	288	0/2/19/27/24/48/36/132	236	2 346 495
2	http://de.wikipedia.org/wiki/Hommingberger_Gepardenforelle	0,9 %	5	248	233	0/0/6/14/17/15/4/177	151	35
3	www.funana.de/hommingberger-gepardenforelle-news.html	2,9 %	5	213	141	0/1/7/15/15/8/1/94	94	2 392 734
4	www.ranking-check.de/gepardenforellen/_hommingberger-forelle_.html	1,8 %	6	303	235	0/0/0/0/0/0/235	159	48 417
5	www.danielgal.de/hommingberger-gepardenforelle/hommingberger-gepardenforelle.html	3,8 %	5	209	147	0/0/16/24/20/19/11/57	126	-
6	http://hommingberger.auctionant.de	1,4 %	5	168	118	0/0/4/8/11/14/2/79	99	-
7	http://gepardenforelle.terrikay.de	4,8 %	5	108	70	0/1/5/14/5/6/1/38	63	-
8	www.kde.cs.uni-kassel.de/lehre/ss2005/googlespam	2,6 %	5	218	139	0/3/18/14/16/10/4/74	103	93 339
9	http://hommingberger-gepardenforelle.malkusch.de	3,7 %	5	187	119	0/0/11/25/22/13/4/44	104	1 344 791
10	http://hommingbergergepardenforelle.blogspot.com		5	88	51	0/0/0/7/3/1/1/39	26	30
MSN								
1	www.hommingberger2gepardenforelle.de/?Hommingberger%20Gepardenforelle	14,1 %	0	194	48	0/0/0/0/0/0/1/47	4	-
2	http://hommingberger-gepardenforelle.malkusch.de	3,7 %	5	187	119	0/0/11/25/22/13/4/44	104	1 344 791
3	www.hommingberg-gepardenforelle.de	6,0 %	4	25	16	0/0/0/1/2/1/0/12	10	-
4	www.hommingberger-gepardenforelle-0.info/_Hommingberger_Gepardenforelle_.html	2,6 %	0	414	224	0/0/0/0/0/0/1/223	167	-
5	www.hbfg.de	1,5 %	3	45	28	0/0/2/4/0/2/1/19	19	-
6	www.hommingberger-gepardenforelle.de	6,6 %	6	882	613	0/6/11/25/62/56/39/414	498	794 741
7	www.hommingberger0gepardenforelle.de/?Hommingberger%20Gepardenforelle	10,5 %	0	204	48	0/0/0/0/0/0/1/47	5	-
8	http://gepardenforelle.nophia.de	4,1 %	5	121	64	0/0/4/18/6/2/3/31	46	2 353 095
9	http://gerech.net	4,4 %	4	37	23	0/0/4/4/2/1/1/11	19	-
10	www.estivo.de/?hommingberger-gepardenforelle	1,2 %	2	158	91	0/0/0/0/1/0/2/88	30	-
Seekport								
1	www.heise.de/ct/SEO-Wettbewerb	1,5 %	6	618	378	0/7/17/54/54/38/29/179	324	1 928
2	www.hommingberger-gepardenforelle.de	4,4 %	6	882	613	0/6/11/25/62/56/39/414	498	794 741
3	http://de.wikipedia.org/wiki/Hommingberger_Gepardenforelle	0,9 %	5	248	233	0/0/6/14/17/15/4/177	151	35
4	www.gepardenforelle-hommingberger-gepardenforelle.de	3,0 %	4	655	400	0/0/3/9/15/20/22/331	328	3 707 863
5	www.hommingberger-gepardenforelle-fisch.de	3,6 %	4	59	46	0/0/2/7/8/3/0/26	29	2 965 088
6	http://gepardenforelle.nophia.de	4,1 %	5	121	64	0/0/4/18/6/2/3/31	46	2 353 095
7	http://hommingbergergepardenforelle.blogspot.com	2,2 %	5	88	51	0/0/0/7/3/1/1/39	26	30
8	www.funana.de/hommingberger-gepardenforelle-news.html	2,6 %	5	208	142	0/2/6/6/16/3/0/109	89	2 392 734
9	www.lenz-online.de/hommingberger-gepardenforelle	3,1 %	4	37	27	0/0/0/0/0/0/0/27	19	-
10	www.hommingberger-gepardenforelle.net	2,5 %	6	860	530	0/3/19/34/32/57/40/345	441	2 088 380

schlüsselt einige wichtige Einflussfaktoren auf. Eine vollständige Analyse aller für eine Top-Position entscheidenden Parameter kann dieser Artikel aus Platzgründen aber nicht leisten. Sie wäre ohnehin nur eine Momentaufnahme mit kurzer Halbwertszeit, da die Suchmaschinenbetreiber ihre Ranking-Algorithmen permanent weiterentwickeln, um Manipulationen einen Riegel vorzuschieben.

Wer sich durch den Wettbewerb ermutigt sieht, seine Seiten

besser für Suchmaschinen anzupassen: Das Suchmaschinen-Optimier-Einmaleins haben wir in [1] beschrieben. Um immer über die neuesten Entwicklungen auf dem Laufenden zu sein, empfiehlt sich ein Blick in einschlägige Foren [8, 9, 10]. (jo)

Literatur

- [1] Jo Bager, Aufsteiger, Websites mit dem Internet Business Promoter suchmaschinengerecht aufbereiten, c't 9/05, S. 158
- [2] Jo Bager, Turbolift ins Ungewisse,

Manipulationsmaschen bei Suchmaschinen, c't 9/05, S. 164

- [3] Jo Bager, Jan Becker-Fochler, Forellenschwemme, Wettbewerb zur Suchmaschinenoptimierung: Ergebnisse des ersten Stichtags, c't 12/05, S. 38
- [4] Homepage des Wettbewerbs mit detaillierten Ergebnissen, www.heise.de/ct/SEO-Wettbewerb
- [5] Listen der Top-100-Treffer, www.vwp-online.de/hommingberger-gepardenforelle
- [6] Alexa-Homepage mit Traffic Rankings, www.alexa.com

[7] SCigen – An Automatic CS Paper Generator, <http://pdos.csail.mit.edu/scigen>

- [8] SEO-Forum bei Abakus Internet Marketing, www.abakus-internet-marketing.de/foren
- [9] SEO-Forum bei Suchmaschinen-Tricks, www.suchmaschinentricks.de/forum
- [10] SEO-Forum und News bei WebmasterWorld.com, www.webmasterworld.com

