Implications of search engine spam on the visibility of South African e-commerce Web sites

F. Mbikiwa
e-Innovation Academy
Cape Peninsula University of Technology
Cape Town, South Africa
ferniembikiwa@gmail.com

M. Weideman
e-Innovation Academy
Cape Peninsula University of Technology
Cape Town, South Africa
meliusw@yahoo.com

Contents

Key words: Spam, search engines, e-commerce Web sites, search engine optimization

---

1 Introduction

According to Peng, Trappey and Liu (2005:476), the Internet has made it essential for e-traders to adopt e-commerce as a means of obtaining a competitive advantage over other firms. Darch and Lucas (2002:148) indicate that e-commerce has changed business processes that were evident in traditional businesses.

However, e-commerce Web sites should attract visitors, and one way of accessing these Web sites is through the use of search engines (Thelwall 2000:150). Ranking high in search engine results can be considered a competitive strategy. According to a study done by Nielsen (2004), 88% of users start with search engines when looking for information on the Web. This has resulted in Web site owners adopting certain practices to ensure that their Web sites gain top rankings in search engine results. Fetterly, Manasse and Najork (2004) are of the opinion that high placement in search engines is one of the strongest contributors to a commercial Web site's success.

Machill, Neuberger and Schindler (2003:54) state that some Web site designers often adopt practices to assist the Web site in achieving high rankings while Henzinger, Motwani and Silverstein (2002) substantiate this argument by stating that some Web designers deliberately manipulate their rankings on search engines. It is not clear how search engines view search engine optimization practices.

According to Sullivan (2001), most search engine policies are vague as to what is considered search engine spam. Perkins (2001) has defined spam as 'any attempt to deceive search engine relevancy rankings'. This type of spam is different from the traditional e-mail spam, which refers to sending bulk messages to different e-mail addresses without the consent of the e-mail account holder (Adam 2002:91).

Some research has been done regarding what is considered as spam. However, no evidence could be found of empirical work, which has included search engine spam and its effects on Web site visibility. The aim of this study was to determine the implications, if any, that spam has on e-commerce Web sites with respect to search engine indexing and rankings. Most search engine policies state that Web sites containing spam will not be indexed.

A variety of search engines in various forms are discussed in this article, followed by other research done on search engine optimization (SEO). Search engine spam is a vague concept, interpreted in different ways by different authors, Web site developers and search engine designers. The authors therefore inspected 11 different types of spam. The methodology involved inspection of a number of e-commerce Web sites and comparison with search engine exclusion policies. It was found that only about 20% from the sample of South African e-commerce Web sites contained some form of search engine spam.

---

2 Search tools

Search engines and directories are considered as two of the most important tools for locating/retrieving information (Thelwall 2002:101) and are also seen as the primary tools for users to find Web sites (Sullivan 2002a). At the same time, however, many users seem to experience problems in the use of search engines for retrieving relevant information (Weideman 2005:12). According to Sullivan (2002c), search tools can be categorized as crawler-based search engines, directories or meta search engines. Crawler-based search engines create their

listings automatically (Sullivan 2002b). Green (2000:125) describes a directory as 'a predefined list of Web sites compiled by human editors and categorized according to a certain topic or subject'. The same author claims that Web sites that are listed in a directory are likely to stay there longer than Web sites that have been indexed by a search engine, due to the manual process involved in compiling a directory. According to Thurow (2003:29), top directory listings are based on the following criteria:

- Directory category
- Web site's title
- Web site's description.

When dealing with meta search engines, the searcher is not querying the database of the meta search engine but is actually searching across several search engines and Web directories (Green 2000:127). Barker (2005) describes how meta search engines work by explaining that once a searcher submits a keyword to a meta search engine search box, the search query is transmitted simultaneously to several search engines and their databases. This author also states that meta search engines do not own any database of Web pages but rather send the search terms to the databases maintained by search engine companies.

Some contradiction is apparent in the literature on search engines. On the one hand it is claimed that large numbers of Internet users rely on search engines to find relevant information. At the same time it is also claimed that users experience trouble in finding relevant information on the Internet. It would seem to imply that there is a lot of frustration among Internet users when it comes to quickly finding relevant information through the use of search engines.

The search tools that receive the most traffic according to Nielsen/Netratings (2005) are reflected in Table 1.

Table 1 Top search tools in 2005 (Nielsen/Netratings 2005).

Position	Provider	June 05	July 05
1	Google Search	47,0%	46,2%
2	Yahoo! Search	22,3%	22,5%
3	MSN Search	12,5%	12,6%
4	AOL Search	5,5%	5,4%
5	My Way Search	1,8%	2,2%
6	AskJeeves Search	1,8%	1,6%
7	Netscape Search	0,9%	1,6%
8	Dogpile.com Search	0,8%	0,9%
9	iWon Search	1,0%	0,9%
10	EarthLink Search	0,8%	0,8%

3 Search engine optimization (SEO)

SEO aims at improving the ranking of a Web site on search engine result lists (Wikipedia.com 2005b). SEO can be regarded as a mechanism for drawing potential customers (Ngindana 2005:41). This has culminated in many Web designers implementing search engine optimizers in an attempt to heighten the retrievability of their Web sites by search engines (Machill *et al.* 2003:53).

According to Zhang and Dimitroff (2004:666), search engine optimization is a process applied to Web pages to enhance access to the Web page by search engines. The authors are further of the opinion that SEO has become a very complex and sophisticated practice that requires constant research, practice and re-evaluation to be effective. Google (2005) argues that certain dubious SEO practitioners have participated in unethical marketing practices that have the sole purpose of influencing search engine results.

Fetterly *et al.* (2004) point out that there are some SEO firms that guarantee high placements for Web sites in search engine result lists. The Web sites are loaded with keywords in an attempt to appear relevant to search engines. This is classified as spam. Some SEO practices have in the past raised a red flag in terms of ethical issues. According to Weideman (2004:910), there are two main ethical questions that arise regarding some SEO practices:

- A given Web page is excluded from search engine rankings while it deserves to be there due to high quality or relevance of content
- A given Web page is included in rankings resulting from payment by the owner, regardless of the quality or relevance of its content.

It appears as if SEO is a complex, easily misunderstood process. A Web site designer, aiming to remain on the right side of the search engine law by optimizing a Web page according to some very specific guidelines, might find that that very Web page is accepted by one but banned by another search engine without apparent reason.

4 Search engine spam

Traditional e-mail spam is a serious problem for most users today, and has spawned a large number of specialized software programs to combat it (Greiner 2006:12). However, search engine spam is defined differently. According to Wilkinson (2004), search engine spam can be defined as anything that constitutes unethical practices within SEO, and this includes manipulating search engine spiders and redirecting users to inappropriate content. The artificial boosting of Web site rankings reduces the quality of search results, resulting in the quality of these search results becoming questionable (Anon 2002). However, Sullivan (2001) argues that there is no set definition for what is considered to be spam. Sullivan cites an example of cloaking, which is considered as spam by Google but in contrast both AltaVista and Inktomi allow it under certain circumstances. Thurow (2003:218) regards search engine spamming as a waste of time. This author states that the amount of money and time spent on using techniques that are considered as spam can be better spent on optimizing Web sites using the correct techniques that will please both the user and search engines. Research conducted by different authors has identified the practices below as spam.

Search engine spam appears to be a minefield of potential problems, which require a very skilled designer to navigate successfully. Even highly rated authors imply that some search engines employ methods of spam that they themselves criticize! One solution appears to be to aim to please everybody – apply optimization to a Web page and attempt to remain on the safe side in terms of the specifications of most search engines. This could imply using the strictest limit in each case. For example, if 10 search engines specify the maximum number of keywords in the keyword metatag to be between 50 and 100, with nine of them listing 100 as the maximum, one should use only 50.

4.1 Cloaking

Rowlett (2003) has defined cloaking as a process where visitors to a Web site are shown a different page to what the search engine spiders have seen. According to Sullivan (2003), few issues have been as controversial as cloaking. This author argues that even though most search

engines have guidelines against cloaking, some still allow it, while others go as far as practising it themselves. While Thurow (2003:227) considers cloaking to be spam, Sullivan (2003) argues that cloaking should not be considered as spam because it does nothing to satisfy a search engine's algorithm – it is just a way of ensuring that targeted content is delivered.

4.2 Doorway pages

A doorway page is a Web page which redirects a human visitor to a different, more human friendly Web page than a search engine crawler. Thurow (2003:227) argues that the primary reason why Web designers create doorway pages is to obtain high rankings with search engines. A doorway page has also been defined as a Web page that has been designed to deceive search engines into ranking it higher for one or more particular keywords.

4.3 Invisible text

For Web sites to receive high rankings, keywords contained in search queries must be present in a Web page. Some Web designers code keywords that are the same colour as the background of a Web page on the Web page itself, resulting in the keywords being invisible to the human reader. Similarly, text in the smallest font possible (called tiny text) is readable by crawlers but not humans. These keywords are intended to be visible to spiders, and not to users that visit the site (Anon 2002; Thurow 2003:222). Other implementation methods of invisible text include hiding text behind layers of graphics and placing text at the bottom of oversized pages (Dunn 2004). Invisible text is even used inside meta tags (Chambers 2005:57).

4.4 Artificial link farms

Link farms are Web pages containing nothing but thousands of hyperlinks to a variety of Web pages. Henzinger *et al*. (2002) state that the main purpose of artificial link farms is to manipulate systems that use the number of incoming links to a Web site to determine relevancy. Search engines rely on link analysis to determine the relevancy of Web sites to search engines. A study conducted by Sullivan (2002d) indicates that what a Web site links to represents a major component of how that particular Web site is ranked. Sullivan has also stated that the best way to be listed in Google is to build links to the Web site, as crawlers follow links and index new pages (Sullivan 2004). This has resulted in Web designers attempting to artificially increase link popularity by creating multiple Web sites with the intention of linking the sites to one another, a practice commonly referred to as artificial link farms. However, Thurow (2003:224) points out that search engines consider link farming to be spam, and Web sites that implement artificial link farms are likely to be penalized by search engines.

4.5 Keyword stacking/stuffing

Thurow (2003:221) defines keyword stacking as the repeated use of a keyword or keyword phrase to artificially boost a Web page's relevancy in search engines. Thurow states that keyword stacking and keyword stuffing often have the same meaning and therefore the two terms are used interchangeably in this article. According to Wilkinson (2004), even though keyword stacking may help a Web site's keyword relevancy attain a higher rank, it is likely that the high ranking will not last long and also likely that the Web site may be penalized by search engines. The same author states that keyword stuffing also refers to loading the page with content that extends so far down the page that it is unlikely anyone will continue to scroll down. This technique is likely to be used by designers who rely on splash pages for their index pages.

4.6 Hidden links

Hidden links are hyperlinks on a Web page, which are visible only to crawlers, not to human visitors. Google, which is one of the top search engines (Nielsen/Netratings 2005), places emphasis on the number of inbound links when determining relevancy (Wikipedia 2005a). This

has resulted in Web designers creating these hidden links (Thurow 2003:223). According to Anon (2005), hidden links are commonly used to increase the link count of Web pages.

4.7 Page redirects

A page redirect involves the placing of HTML code that will redirect a site visitor from a page that is designed only for a search engine to another page designed for human visitors (Anon 2002; Thurow 2003:225). According to Thurow, spammers create Web pages that are optimized with certain keyword phrases. These pages are then submitted to search engines, but whenever a searcher clicks on this Web page they are automatically redirected to another page, known as a destination page.

4.8 Duplicate pages

This process involves spammers slightly modifying Web pages and resubmitting them to search engines. This often results in Web pages with identical content but different title tags appearing in search results (Anon 2002; Thurow 2003: 226). According to Dunn (2004), Webmasters create a Web page, and then duplicate Web pages of that particular site are created. Each page is often optimized differently from the other to get multiple placements in search engines.

4.9 Domain spam

Domain spam refers to the purchasing of several domain names and creating Web sites with identical content. This is done for the sole purpose of achieving multiple listings in directories with the intention of achieving link popularity and more traffic. Web designers anticipate that the link popularity will improve search engine rankings (Thurow 2003:227).

4.10 Automatic submission

This process involves automatically submitting Web sites to search engine directories, using software. This process is often done repeatedly, causing the search engines to be flooded with similar Web sites (Anon 2004; Dunn 2004). Anon (2004) also states that automatic submission is done by visiting an 'auto-submit URL' site, which then submits the URL to many search engines that are in its database. Dunn (2004) points out that automatic submission floods the bandwidth of search engines as Web sites are often submitted repeatedly within the same search engine.

4.11 Meta tag stuffing

Meta tag stuffing refers to the practice of repeating keywords in the meta tags and using keywords which are unrelated to the site's content (Wikipedia.com 2005c). Thurow (2003) states that search engines initially used meta tags to determine the relevancy of Web sites to search queries entered by users. Furthermore, owing to meta tag stuffing, most search engines do not use them for ranking purposes. A study done by Wallace (2003) indicates that some unethical Web designers place high traffic keywords that are not related to the Web page in meta tags to generate more traffic.

5 Methodology

A number of methods were considered for this research project, as listed below.

5.1 Non contact methods

5.1.1 Postal questionnaire
Some of the problems associated with this methodology include non-response, questionable validity and reliability, respondent bias and lack of control.

The non-response and lack of control problems disqualified postal questionnaires as a research method. Similarly, the potential validity problems indicated that this method could yield invalid results. As the proposed model cannot be based on unreliable data, this method was disqualified.

5.1.2 Internet questionnaire
An Internet-based questionnaire has the potential of delivering a high number of respondents, owing to the high number of Internet users. However, it was scrapped since it would be impossible to establish the environment of participants.

5.1.3 Case study
It is said that a case study should be carried out on the various attributes of one individual. Some authors describe a variety of case studies, all concentrating on the behaviour of one small group of humans, while others claimed that the case study method is useful to describe a unit rather than for the testing of a hypothesis. It was considered that the size of the sample in the case study method would be inadequate.

5.2 Contact methods

5.2.1 Personal interview
Ample evidence exists that personal interviewing is an expensive research method, which ruled it out for this project.

5.2.2 Telephonic interview
This method suffers from the similar drawbacks as personal interviews. Elements such as the cost factor, time, lost production, interviewer training and the fact that this work is very tiring were mentioned.

5.2.3 Laboratory experiment
Exorbitant costs would make this method unsuitable to implement.

5.2.4 Field experiment
As far back as 1970, it was claimed that 'a carefully controlled and rigorously conducted experiment is probably the most exact and refined method of research' (Lovell and Lawson 1970:82). Smit stated that field experiments provide the element of control expected from scientific research (Smit 1983:46). It was noted that some of its advantages include lower cost (since one person can administer it), and full control over the situation (since the administrator is present at every session) (Huysamen 1994:150,151).

The authors decided to use the field experimental method, based on Smit's mention of the element of control, supporting the scientific nature of the proposed fieldwork.

The methodology to execute this research project included the following steps:

- A checklist of what is generally considered as spam was compiled from the literature, as well as a search engine exclusion policy analysis.
- The policies of five search engines namely Google, Yahoo!, Ananzi, AksJeeves and Altavista were examined to determine what they consider as spam.
- Forty-seven active e-commerce Web sites were selected as described below.
- These Web sites were checked against keyword verification software to determine if they are indexed by the above-mentioned search engines.
- The Web sites were then checked against the spam checklist to determine if they contain any of the practices that were identified as spam.

Over 4000 e-commerce Web sites were obtained from the Cape Chamber of Commerce and onlineshopping.co.za. A total of 47 Web sites were randomly selected from this population. This selection was based on the allocation of a random number to each Web site (using a spreadsheet's randomize function). These Web sites were then inspected, using dedicated software, and the results were tabulated to indicate with which search engines they are registered. The resultant list also reflected which Web site was registered with which search engine as well as Web sites that were not indexed by any search engine. In the next phase, various spam elements were identified and checked for in these 47 Web sites. The following process was used to detect the occurrence of spam:

- Keyword stuffing. The meta tags of the specific pages were manually inspected to see if they contained any keyword stuffing.
- Invisible/tiny text. These were checked by highlighting the page and dragging the cursor from the top to the bottom of the page. Any invisible text would be highlighted.
- Meta tag stuffing. Meta tags were manually inspected for repeated keywords or keywords that were not related to the Web site.
- Page redirects. This was identified by viewing the URL in the status bar before clicking a link. By looking at the status bar before clicking a link, Java redirect code in the link would be detected. The URL was observed to determine if it changed to another URL before opening the Web page.
- URL spam. The URL was examined to detect if it is long and contains sequences of spam terms.

6 Results

The analysis of the 47 Web sites returned two major findings (Table 2 and Table 3). In the first instance a small number of the analysed Web sites contained spam, and in the second instance there were only two types of spam identified namely 'keyword stuffing' and 'page redirects'.

Table 2 Percentage of Web sites with spam

Total number of analysed Web sites	Total number of Web sites with spam	Percentage of Web sites with spam
47	10	21,3%

Table 3 Web sites used and their presence in search engines

	Random number	Google	Yahoo	AskJeeves	Altavista	Ananzi
www.pcshopping.co.za	0.529810049	Y	N	Y	N	Y
www.ngrcomputers.co.za	0.590298186	Y	Y	Y	Y	Y
www.magafters.com	0.99392923	Y	Y	Y	Y	Y
www.sarugby.com	0.764054842	Y	Y	Y	Y	Y
www.prohampers.co.za	0.7634205	Y	Y	Y	Y	Y
www.intersoft.co.za	0.114912137	N	N	N	N	Y
www.orionsbelt.co.za	0.139283534	Y	N	Y	N	Y
www.mrmattress.co.za	0.934011603	Y	Y	Y	Y	Y
www.soholink.co.za	0.781562232	Y	N	Y	N	N
www.boardgames.co.za	0.062169864	N	N	Y	N	Y
www.gadgetshouse.co.za	0.544735856	Y	Y	N	Y	Y
www.enigmatek.co.za	0.433696508	Y	N	Y	N	Y
www.sabshop.com	0.238287585	Y	Y	N	Y	Y
www.coffee.co.za	0.60258904	N	N	Y	N	Y

www.energyzone.co.za	0.53613596	Y	Y	Y	Y	Y
www.netflorist.co.za	0.894239366	Y	Y	Y	Y	Y
www.volpes.co.za	0.260735456	Y	Y	Y	Y	Y
www.egos.co.za	0.28816306	Y	Y	Y	Y	Y
www.mecerdirect.co.za	0.097452076	Y	Y	Y	Y	Y
www.over2u.com	0.92881439	Y	Y	Y	Y	Y
www.waltons.co.za	0.173680632	Y	Y	Y	Y	Y
www.giftlady.net	0.439258971	Y	Y	Y	Y	Y
www.loot.co.za	0.591205369	N	Y	Y	Y	Y
www.good4u.com	0.496423977	Y	Y	Y	Y	Y
www.timeslice.co.za	0.684989828	Y	Y	Y	Y	Y
www.giftware.co.za	0.696107694	N	N	Y	N	Y
www.jump.co.za	0.660878198	Y	Y	Y	Y	N
www.thegadgetshop.co.za	0.655579834	Y	Y	N	Y	Y
www.musica.co.za	0.614544777	Y	Y	Y	Y	Y
www.prodigital.co.za	0.803625213	Y	N	Y	N	Y
www.streetcar.com	0.603359362	Y	Y	Y	Y	Y
www.shopping.imagecorp.co.za	0.583992672	Y	N	Y	N	N
www.rubberstampsa.co.za	0.400811798	Y	N	Y	N	Y
www.premiumwines.co.za	0.647203898	Y	N	N	N	Y
www.healthspa.co.za	0.478814365	Y	N	Y	N	Y
www.team101.co.za	0.306309692	Y	Y	Y	Y	Y
www.progps.co.za	0.942887179	Y	Y	Y	Y	Y
www.wholesaledigital.co.za	0.855486362	Y	Y	Y	Y	Y
www.hammeronguitars.co.za	0.542882004	Y	Y	N	Y	Y
www.underanafricansky.co.za	0.879475772	N	Y	N	Y	N
www.cacell.co.za	0.483114566	Y	Y	Y	Y	Y
www.fotodigital.co.za	0.543047733	Y	Y	Y	Y	Y
www.southafricanfoodshop.com	0.017139978	Y	Y	Y	Y	Y
www.mynewlaptop.co.za	0.028897853	Y	Y	Y	Y	Y
www.sun-e-shop.co.za	0.387111045	Y	Y	Y	Y	Y
www.ascotdirect.co.za	0.567882678	Y	Y	Y	Y	N
www.healthier.co.za	0.564540938	N	N	Y	N	Y
www.cheaper.co.za	0.772055108	N	N	Y	N	Y
www.shopasave.co.za	0.43198867	Y	Y	Y	Y	Y
www.homemark.co.za	0.104288836	Y	Y	Y	Y	Y
www.josejewels.co.za	0.619885242	Y	Y	Y	Y	N

6.1 Keyword stuffing results

Figure 1 indicates the number of Web sites that contained keyword stuffing. In evaluating the results, it is evident that although the five search engines state that Web pages with keyword stuffing will not be indexed, some were indexed by the same search engines, indicating non-compliance of the search engines to their own policies. However, the figure also indicates Web sites that contained keyword stuffing but were not indexed by the search engines.

Figure 1 Keyword stuffing results

Keyword Stuffing
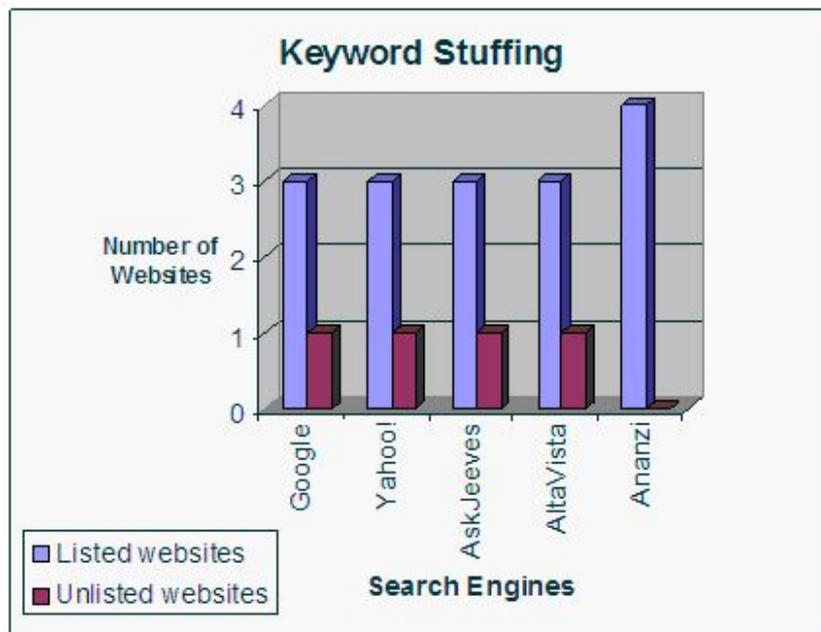
6.2 Page redirects results

During the analysis of page redirects, Google had the highest number of Web sites with page redirects that had been indexed, followed by Ananzi (Figure 2). This implied a possible exclusion of the Web sites by the other three search engines based on the presence of page redirects.

Figure 2 Page redirects results

Page Redirects
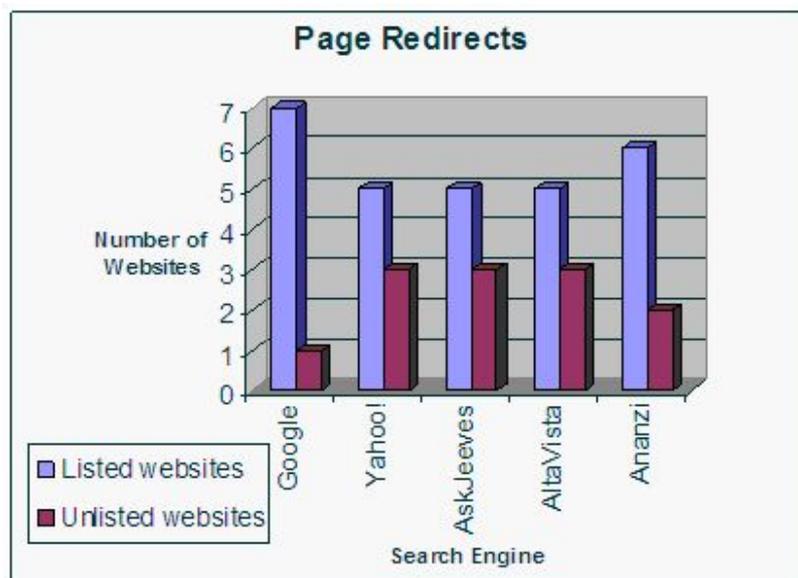
6.3 Results of other spam in the analysis

The following types of spam were not evident in any one of the Web sites analysed:

- Invisible text
- Tiny text
- Hidden links
- Artificial link farms
- URL spam

- Doorway pages
- Meta tag stuffing

7 Conclusion

The objective of this research was to determine the implications that search engine spam have on the visibility of e-commerce Web sites. It also aimed at determining whether search engines comply with their policies and exclude Web sites that do contain spam. From the literature, it was claimed that spam is often applied to distort search engine results, and that Web designers apply 'unethical' SEO practices with the sole purpose of achieving high rankings in search engines.

However, from the empirical results, it was evident that a relatively small number of e-commerce Web site designers are using such spamming techniques. Of the e-commerce Web sites that were analysed, only 21,3% contained spam.

This apparent contradiction could be ascribed to a number of factors. The Web sites inspected in this study were all South African e-commerce ventures. It could be that South African designers are not as acutely aware of the possible ways of spamming search engines. Furthermore, competition in the online world is possibly not quite as strong in South Africa compared to the American market, and it has not become necessary to compete for high search engine result placings at any cost, as it has in the USA and the UK.

Finally, previous research has shown that the South African market tends to follow overseas markets in terms of general trends. If this fact were also true for search engine marketing, local Web designers would be well advised to refrain from using the spam techniques discussed in this article. Standard good practice for Web page design should be followed, while considering the known search engine exclusion policies and algorithms.

8 References

Adam, R. 2002. Is e-mail addictive? *Aslib Proceedings* 54(2):85-94.

Anon, 2002. *SEO code of ethics.* [Online]. Available WWW: http://www.searchengineethics.com/ (Accessed 3 July 2005).

Anon, 2005. *Hidden links and text.* [Online]. Available WWW: http://www.aim-pro.com/helpfiles/hiddenlinks.html (Accessed 7 September 2005).

Barker, J. 2005. *Meta-search engines.* [Online]. Available WWW: http://www.lib.berkeley.edu/Teachinglib/Guides/Internet/metaserach.html (Accessed 3 September 2005).

Chambers, R. 2005. Search engine strategies: a model to improve Web site visibility for SMME Web sites. Unpublished M.Tech thesis, Cape Peninsula University of Technology, Cape Town.

Darch, H. and Lucas, T. 2002. Training as an e-commerce enabler. *Journal of Workplace Learning* 14(4):148-155.

Dunn, R. 2004. *The top 10 worst SEO tactics.* [Online]. Available WWW: http://www.stepforth.com/ (Accessed 18 September 2005).

Fetterly, D., Manasse, M. and Najork, M. 2004. Spam, damn spam, and statistics. Using statistical analysis to locate spam Web pages. In: *Proceedings of the Seventh International Workshop on the Web and Databases,* Paris, 17-18 June.

Google. 2005. *Google information for Webmasters: Webmaster guidelines.* [Online]. Available WWW: http://www.google.co.za/Webmasters/index.html (Accessed 01 March 2005).

Green, D. 2000. The evolution of Web searching. *Online Information Review* 24(2):124-137.

Greiner, L. 2006. Enemy, thy name is spam. *Computing Canada* January, 12-13.

Henzinger, M.K., Motwani, R. and Silverstein, C. 2002. *Challenges in Web search engines.* [Online]. Available WWW: http://www.acm.org/sigir/forum/f2002/henzinger.pdf (Accessed 12 August 2005).

Huysamen, G.K. 1994. *Methodology for the social and behavioural sciences.* Midrand: Southern.

Lovell, K. and Lawson, K.S. 1970. *Understanding educational research.* London: University of London Press.

Machill, M., Neuberger, C. and Schindler, F. 2003. Transparency on the net: functions and deficiencies of Internet search engines. *Info – The Journal of Policy, Regulation and Strategy for Telecommunications* 5(1):52-74. [Online]. Available WWW: http:/www.emeraldinsight.com/1463-6697.htm (Accessed 2 March 2005).

Ngindana, M.W. 2005. Visibility of e-commerce Web sites to search engines: A comparison between text-based and graphic-based hyperlinks. Unpublished M.Tech thesis, Cape Peninsula University of Technology, Cape Town.

Nielsen, J. 2004. *When search engines become answer engines.* [Online]. Available WWW: http://www.useit.com/alertbox/20040816.html (Accessed 14 August 2005).

Nielsen/Netratings. 2005. *Nielsen/Netratings releases top 10 search engine share rankings for July 2005.* [Online]. Available WWW: http://www.netratings.com/pr/pr_050824.pdf (Accessed 29 October 2005).

Peng, Y., Trappey, C.A. and Liu, N. 2005. Internet and e-commerce adoption by the Taiwan semiconductor industry. *Industrial Management & Data Systems* 105(4):476-490.

Perkins, A. 2001. *The classification of search engine spam.* [Online]. Available WWW: http://www.silverdisc.co.uk/articles/spam-classification/ (Accessed 12 September 2005).

Rowlett, D. 2003. *Stop search engine spam!* [Online]. Available WWW: http://www.internetmarketingWebsites.com/spam-review.htm (Accessed 12 June 2005).

Smit, G.J. 1983. *Navorsingsmetodes in die gedragswetenskappe.* Pretoria: Haum.

Sullivan, D. 2001. *Desperately seeking search engine marketing standards.* [Online]. Available WWW: http://searchenginewatch.com/sereport/article.php/2164371 (Accessed 5 September 2005).

Sullivan, D. 2002a. *How search engines work*. [Online]. Available WWW: http://www.searchenginewatch.com/Webmasters.php/34751_2168031 (Accessed 25 August 2005).

Sullivan, D. 2002b. *Intro to search engine optimization*. [Online]. Available WWW: http://www.searchenginewatch.com/Webmasters/print.php/34751_2167921 (Accessed 8 March 2005).

Sullivan, D. 2002c. *Search engine features for Webmasters.* [Online]. Available WWW: http://www.searchenginewatch.com/Webmasters/34751_2167891 (Accessed 25 August 2005).

Sullivan, D. 2002d. *Search engine link popularity.* [Online]. Available WWW: http://www.searchenginewatch.com/searchday/print.php/34711_2159711 (Accessed 3 September 2005).

Sullivan, D. 2003. *Ending the debate over cloaking*. [Online]. Available WWW: http://www.searchenginewatch.com/sereport/print.php/34721_2165321 (Accessed 8 March 2005).

Sullivan, D. 2004. *Submitting to crawlers: Google, Yahoo, Ask/Teoma & Microsoft's MSN.* [Online]. Available WWW: http://searchenginewatch.com/Webmasters/print.php/34751_2167871 (Accessed 28 June 2005).

Thellwall, M. 2000. Commercial Web sites: lost in cyber space? *Internet Research: Electronic Networking Applications and Policy* 10(2):150-159.

Thelwall, M. 2002. Subject gateway sites and search engine ranking. *Online information Review* 26(2):101-107.

Thurow, S. 2003. *Search engine visibility.* Indianapolis: New Riders Publishing.

Wallace, D. 2003. *Spamming techniques that you will want to avoid.* [Online]. Available WWW: http://www.searchrank.com/resources/art003.htm (Accessed 10 November 2005).

Weideman, M. 2005. FOIOTI: An implementation of the conceptualist approach to Internet searching. *South African Journal of Libraries and Information Science*, 71(1):11-25.

Weideman, M. 2004. Ethical issues on content distribution to digital consumers via paid placement as opposed to Website visibility in search engine results. In: *Proceedings of the Seventh International Conference ETHICOMP 2004,* 14-16 April 2004, University of the Aegean, Syros (2):904-915.

Wikipedia.com. 2005a. *Search engine*. [Online]. Available WWW: http://en.wikipedia.org/wiki/search_engine (Accessed 29 September 2005).

Wikipedia.com. 2005b. *Search engine optimization*. [Online]. Available WWW: http://en.wikepedia.org/wiki/serach_engine_optimization (Accessed 29 September 2005).

Wikipedia.com. 2005c. *Spamdexing*. [Online]. Available WWW: http://en.wikipedia.org/wiki/Spamdexing (Accessed 29 September 2005).

Wilkinson, T.A. 2004. *Just say no to SEO spam*. [Online]. Available WWW: http://www.w-edge.com/ (Accessed 23 August 2005).

Zhang, I. and Dimitroff, A. 2004. *The impact of Webpage content characteristics on Webpage visibility in search engines (Part I).* Information Processing & Management 41(2005): 665-690 [Online]. Available WWW: http://jis.sagepub.com/cgi/content/refs/30/4/310 (Accessed 12 February 2005).

Disclaimer

Articles published in SAJIM are the opinions of the authors and do not necessarily reflect the opinion of the Editor, Board, Publisher, Webmaster or the Rand Afrikaans University. The user hereby waives any claim he/she/they may have or acquire against the publisher, its suppliers, licensees and sub licensees and indemnifies all said persons from any claims, lawsuits, proceedings, costs, special, incidental, consequential or indirect damages, including damages for loss of profits, loss of business or downtime arising out of or relating to the user's use of the Website.

top