Application of best practice towards improving Web site visibility to search engines: a pilot study

M. Weideman  
Faculty of Business Informatics  
Cape Peninsula University of Technology  
Cape Town  
South Africa  
meliusw@yahoo.com

R. Chambers  
Faculty of Business Informatics  
Cape Peninsula University of Technology  
Cape Town  
South Africa  
rickardchambers@yahoo.com

Contents

1. Introduction  
2. Web site visibility status quo  
3. Implementation  
4. Results  
5. Conclusion  
6. Acknowledgements  
7. References  
8. Appendix A  
9. Appendix B

Key words: Web site visibility, meta tags, search engine, frames

1 Introduction

Many hours can be spent on creating a Web site before finally introducing it to the world. The content could be informative and presented logically. The design could be outstanding and the user interface functional. Even the unique customers, purpose of the site, simple
navigational strategies and all the essential aspects of Web design could have been considered during implementation. In many developers' eyes, such a Web page could be seen as flawless. The hard reality, though, is that a well-designed site will not be successful if the potential customers cannot find the site.

Often a developer submits Web pages to the search engines and then wants to move on to other work, under the impression that the task is complete. In most cases, however, the developer's responsibility should also include constantly improving the visibility (ability to be seen or capability of being readily noticed or referred to) of the Web site by having the pages indexed and ranked high in the search engines, and to keep them there. Exceptions include those cases where it is not crucial that a Web site be visible (e.g. personal Web pages, pages where paid methods are used to achieve high rankings, etc). This way, the developer assures constant Web site visibility that, in turn, should generate constant traffic to the Web site.

In practice, searchers would like to see the item that they are searching for on the first result screen, and in the top ten hits. Businesses would also like their Web page to appear in this set, where it is most likely to capture the interest of the searcher (Rowley 2001:203-204). Therefore, Web developers should invest in a search engine strategy that is based on the effective use of techniques and programming tools. Examples include meta tags or ALT tags to increase the chances of the Web site being listed in high positions in search queries. One author states: 'A sound search engine strategy provides mechanisms for allowing the frequent evaluation of the site visibility' (Constantinides 2002:2633).

However, no empirical evidence could be found in the literature of a study that summarized the elements of such a strategy, or one which provided a pilot study with a proposed solution to this problem. The objective of this research was therefore to provide a Web site visibility best practice proposal, coupled with a pilot study containing an implementation of these elements.

The pilot study focused on a live, active Web site provided by David Cowie, author of www.sa-cycling.com.

Figure 2 in Appendix A is a screenshot of this Web site before any changes were made, and Figure 3 is a screenshot after the redesigning process. Detailed steps follow, explaining the changes that were implemented to make the site more visible to most search engines.

Step 1: The first step was to use an evaluation program to determine the chosen Web site's visibility, current ranking, key word usage and other relevant features. WebPosition Gold's 'Critic' was used for this purpose because it provided an up-to-date analysis of the Web page for most search engines. WebPosition Gold's Page Critic also contains a knowledge base of recommended optimization parameters, which are updated monthly for each of the major search engines. Table 2 represents the list of WebPosition Gold's Page Critic recommendations based on statistical averages for top ranking pages and the proposed action to be taken.

Step 2: The results gathered from WebPosition Gold were recorded, inspected and analysed, and the reasons (if any) for low ranking, or any other situation which could be improved to deliver more hits, were determined.

Step 3: The Web site was then monitored on a daily basis with the WebPosition Gold program for a number of weeks, to establish a trend, which provided valuable information to track ongoing progress. By establishing a trend beforehand, one could see exactly how the site ranking could change over time for each key word and search engine.
2 Web site visibility status quo

One author states that search engine ranking is an important feature: 'Many times, a search marketer has received that dreaded call from the boss, complaining that when the company name or trademark is typed into a search engine, the first thing that appears on the results page is their #1 competitor' (Hursh 2004).

Most often Web site owners will simply launch their site, do nothing to ensure visibility and wait for surfer traffic that could never come. According to Guenther, it is an area where project teams spend the least amount of time and thought (Guenther 2004a:47–48). It could be compared to buying an expensive user-friendly phone system and then having an unlisted phone number, or placing an attention-grabbing billboard in the middle of a forest next to a dirt road (Subia Creative 2002b).

A potential buyer starts a search for information with the hope of finding a solution to current needs. Those Web sites displayed on the first page of the returned results will enjoy a greater market share of search engine referred leads. The pages in the highest ranked positions will receive the greatest share. Thelwall states: 'Simple, mainly non-technical tests need to be carried out on any newly designed Website, including checking how easy it will be for potential customers to find it, by typing relevant key words into major search engines' (Thelwall 2000:150).

2.1 Search services

There are two main types of search services, namely directories and search engines. Yahoo, MSN and Looksmart are referred to as directories. Directories use human editors to create their listings. When submitting a site to be included in a directory, a human editor reviews the site and determines whether or not to include the site in a specific category in the directory. Search engines on the other hand, such as Google, AltaVista and Intomi, use special software robots called crawlers, to retrieve information from Web pages. These search engine crawlers find and fetch Web pages (a process called crawling or spidering) and construct lists of words and phrases found on each Web page. When an end-user searches for key word phrases found in their index, the search engine tries to match the words typed in a search query with the Web pages in its index that are most likely to have the information (Douglas 2003:45–47; Thurow 2003; Thelwall 2000:152). These two search services can increase the exposure of an online organization into the Web marketplace by allowing online customers to locate and easily access the site (Constantinides 2002:2634).

2.2 Auto submission firms

Some Web site owners invest in search engine submission firms, which would like owners to believe that monthly, multiple, automatic submissions to tens of thousands of search engines will produce a major increase in Web site traffic. Several sources warn Web owners to avoid such gimmicks: 'search engine referred traffic will only result from a well-optimized
Website, that is manually submitted to the major search properties. Avoid auto submission firms who promise otherwise (Subia Creative 2002a). Free auto submission services are often of dubious origin, designed to obtain users' e-mail addresses and sell them to spammers.

At this point it must be noted that the use of the term 'spam' in the context of search engines and Web site visibility differs from the more general concept of receiving unsolicited e-mails. Spam in search engine context refers to dubious techniques employed by Web site designers (e.g. cloaking, key word stuffing, link farms, etc). These techniques all serve the same purpose, namely to fool the search engine algorithms into allocating a higher ranking to the relevant Web page than what it deserves based on its content.

Auto submissions could produce junk mail but no search engine listings. 'In fact, the top 18 search engines account for over 99% of all searches performed on the Web' (Subia Creative 2002a). The others, except for a few industry specific directories and portals, are essentially worthless to Web site traffic (Search engine optimization 2004). Some auto submission firms often advertise large numbers of search engines to which they will submit your site. However, these firms are often no more than a free-for-all site. These sites are simply random lists of links to other sites, which very few people really use, called link farms (Konia 2002:13–14). The purpose of a link farm is to artificially increase the number of in-links to a given (paying) Web page, thereby potentially lifting its ranking, regardless of the value of the contents.

2.3 Visibility techniques

There are several ways of increasing a Web site's visibility to most search engines. Many sources claim that the first and most effective way to make a site more visible is to have it listed with the most popular search engines, as this allows their crawlers to visit and index these pages (Rowley 2001:205–207; Srinivasan, Ruiz and Lam 1996:79–81; Thurow 2003; Nobles and O’Neil 2000; Weideman 2004a). Table 1 lists some of the most popular search engines, each with their unique audiences, pages viewed per person and the average time spent on a page per person.

Clear site objectives should be established prior to the development of the Web site, which provide the necessary foundation from which specific design elements can be identified. Some of these design elements are general and are expected from any site that is well designed (Guenther 2004b:54–56).

Table 1 Popular search services (Dickson and Marshall 2004:34)
Visibility through a search engine requires careful attention to issues such as graphics, frames, metatags, code, content and several other back-end techniques. Although directories do not pay that much attention to metatags, meta description and other back-end code, one should still implement these design standards to make provision for those search services that do still use it (Nobles and O'Neil 2000; Thurow 2003; Thelwall 2000:152; Konia 2002:15–16).

According to Nobles and O'Neil (2000:94), many of the engines are allowing Web developers to submit only the main page of a site. Therefore, one should ensure that the Web site's pages are found by the crawlers. Site maps are tools that assist in increasing the usability of the Web site and would also direct the crawler to the important pages. One author claims that 'site maps are especially relevant since many users do not actually "read" content, but instead skim the information, looking for items of relevance to their purpose' (Trunble 2003:27). The site map consisted of just a simple link at the bottom of the home page, to provide users and search engines with an easy way to get to most of the Web pages. It does not have to cover every page and it can be broken down into multiple categories. According to Dickson and Marshall (2004:28–36), a site map provides assistance not just for search engines, but also for browsers.

**Table 2 WebPosition Gold's Page Critic recommendations**

<table>
<thead>
<tr>
<th>Area description</th>
<th>Current count for SA-Cycling</th>
<th>WebPosition Gold's suggestion</th>
<th>Proposed action</th>
</tr>
</thead>
<tbody>
<tr>
<td>TITLE area keyword frequency</td>
<td>2</td>
<td>1</td>
<td>Decrease key word frequency</td>
</tr>
<tr>
<td>TITLE area word count</td>
<td>4</td>
<td>3 – Jul</td>
<td>Increase TITLE word count</td>
</tr>
<tr>
<td>TITLE area key word prominence</td>
<td>44.40%</td>
<td>53%</td>
<td>Increase key word prominence in TITLE</td>
</tr>
<tr>
<td>Meta key words word count</td>
<td>10</td>
<td>15–37</td>
<td>Add more keywords describing the site</td>
</tr>
<tr>
<td>Meta description area key word</td>
<td>42.00%</td>
<td>62%</td>
<td>Increase key word prominence in meta description</td>
</tr>
<tr>
<td>ALT area key word frequency</td>
<td>0</td>
<td>1–Mar</td>
<td>Increase key words in ALT area</td>
</tr>
<tr>
<td>ALT area word count</td>
<td>0</td>
<td>52–Jan</td>
<td>Increase key words in ALT area</td>
</tr>
<tr>
<td>ALT area key word prominence</td>
<td>0%</td>
<td>59%</td>
<td>Increase key word prominence in ALT area</td>
</tr>
<tr>
<td>Body text key word frequency</td>
<td>0</td>
<td>1–Aug</td>
<td>The key word frequency should be increased if the frames are removed.</td>
</tr>
<tr>
<td>Body text word count</td>
<td>10</td>
<td>257–385</td>
<td>The suggested word count should be implemented if the frames are removed.</td>
</tr>
<tr>
<td>Body text key word prominence</td>
<td>0%</td>
<td>55%</td>
<td>The key word prominence should increase if the frames are removed.</td>
</tr>
<tr>
<td>Whole page word count</td>
<td>6</td>
<td>293–555</td>
<td>The overall page word count would increase if the frames are removed.</td>
</tr>
</tbody>
</table>

3 Implementation

3.1 Logo, TITLE, meta KEYWORD tags and meta DESCRIPTION tags
An earlier study indicated that the value of metatags in general, in improving Web site visibility, seems to be decreasing (Weideman 2004b). However, search engines do seem to reward key words in larger type, bold type and italic type with a higher-ranking value (Thurow 2003). The first focus was on the Web site logo. The logo, an image, consisted out of the South African flag and the text 'sa-cycling' as part of the logo – see Appendix A, Figure 2. The logo was changed to only an image of the South African flag, and the text 'SA-Cycling' was then placed next to the logo in bold. These larger, bold, italic key word amendments were also incorporated in the Web pages' headlines and paragraph headers where appropriate.

The TITLE tag provides the description that appears at the top of the browser window to let an Internet user know where they are. It should be a richly written, key word-dense sentence that accurately describes the content of that particular Web page (Galon 1999). The name of the company/Web site in the TITLE tells the visitor very little about the content of the particular page and does not tell the crawler what type of content the page has. Each and every page of a Web site should have a unique, human generated TITLE. Since none of the major engines are case sensitive, it is considered in order to list key words in lower case only. One should also try to avoid repeating key words more than two or three times and never list the same key word twice in a row. If the same key words are used excessively, it could lead to the search engine algorithm blacklisting a Web page as being guilty of spam.

The TITLE tag should be focused on describing the content of the particular Web page, in a key word-rich manner (Craven 2001:203). It was decided to change the title from 'South African Cycling – http://SA-Cycling.com/' to 'South Africa Cycling – news, fun rides, tours, clubs and much more'.

The meta key word tag was once thought to be the golden key to high search engine rankings but, according to several sources (Craven 2001:203; Konia 2002:15), its importance has dropped dramatically over the last few years. Although most search engines no longer recognize the meta key word tag, it was still thought worthwhile to include it for the smaller search engines that still use it (e.g. Ananzi excludes Web sites without meta tags). It is also wise to add some common spelling mistakes, in case a user miss-spells a word in a search query, for example 'accommodation, acommodation, accomodation, acomodation'.

Key words used by 15 other cycling Web sites with a high ranking were gathered and compiled into the following list of key words, which was added to SA-Cycling's key word section: 'road bikes, roadbikes, cycle, cycles, road cycling, roadcycling, ride, SA cycling, SA-Cycling, sacycling, team, cycle club, cycling club, cycling, South Africa cycling, cycle touring, cycle tour, WP cycling, best cycling tours, fun ride, funrides, cicling, bicycle, bicycling, rapport, giro del capo, giro, cape argus, david cowie, bike touring, RIP, hill, cycle news, cycling news, pedal power, PPA, WPPA, track cycling, rules of cycling, Tour de France, professional cycling, pro cycling, procycling, track, fixed-gear, ride'.

For the meta description tag area (text which usually forms the basis for the description of the site on the search engine results-page), there was little change needed, except for considering the key word prominence (how close to the start of a paragraph the key word appears). According to Konia (2002:18) this is one of the most important factors in determining page relevancy. The original description tag was:

'The South African Cycling Website is provided to cover all topics of Cycling in South Africa: Touring, Important events, Racing, Pedal Power organisations, Local Provincial Bodies, National Federation, Mountain Biking and Club information’ The new version is: 'South African Cycling Website provides cover on all topics of Cycling in South Africa: Touring, Important events, Racing, Pedal Power Organisations, Local Provincial Bodies,
National Federation, Mountain Biking, road deaths and club information'.

The text has been re-written so that it still reads well to the visitor but also positions the key word phrase closer towards the beginning of a sentence or paragraph. The major difference in the new description tag was the removal of the words that were not key words, including words like or, the, and, is, etc. In the process the key word prominence was lifted from 42.9% to 65%.

### 3.2 Graphics, JavaScript and frames

ALT tags, which are text associated with images or graphics, were coded to describe the graphics on the Web page. Search engines do not read graphical elements, and some search engines do consider text used in these ALT tags as key words. ALT tags must accurately describe the particular graphic. An example would be the ALT tag for the logo of the Web site:

```html
<img src="Pics/logo.gif" width="66" height="50" align="top" name="SA-Cycling"
ALT="SA-Cycling Logo"><font size="+7">
```

The crawler can now interpret the name of the graphic (logo.gif) and read the ALT text that describes the image. It can also read the width and height of the graphic, but it cannot interpret the graphic itself.

Search engines encourage the use of ALT tags for several reasons:

- Some users turn graphics off, so that Web pages load faster;
- voice-output screen readers (benefiting the blind) will not 'read' a non-text element'; and
- older browsers cannot reproduce animations, or use applets.

In all three cases, the presence of ALT tags ensures that users have an idea of the graphics image they cannot see, regardless of the reason.

In an interview conducted by Dickson and Marshall (2004:28.36), the interviewee strongly disapproves of the use of JavaScript, Flash and session IDs. The argument is that a Web crawler views a Web site as seen through a text browser. The authors go even further, saying that Google might look for URLs in JavaScript, but states that the average crawler would not be able to do so (Dickson and Marshall 2004:28–36). The success resulting from the removal of JavaScript is highly dependent on a Web designer's personal point of view.

Crawlers will often ignore frame-based sites beyond the home page, omitting them from their searchable databases and significantly reducing the chance of users visiting them (Thelwall 2000:157). AltaVista and other search engines index each frame as a distinct page, which would cause the site to be displayed partially.

These search engines will also index each pane of the frame window as a separate page. This means that if the content matching a query is in a pane, visitors clicking on those links will see only the pane, not the full page as it was originally designed.

It was decided to remove the frames completely and redesign the content as one html page (e.g. 'SA-Cycling – homepage.html' in Appendix A, Figure 2). This page contains all the code for the home page and separate pages for 'news.html, contacts.html' etc. These changes are expected to make SA-Cycling much more visible to most crawlers (Nobles and O'Neil 2000).
3.3 Site map, third-party references, newsletters and e-mail contacts

A site map (a Web page listing of all sections of a Web site) was created to increase the visibility of the Web page, boosting the effectiveness of hyperlinks. Many search engines allow only the submission of the main page of a site, which limits key words and hyperlink elements drastically. By the addition of a site map, the developer not only provides the crawler with an easy navigation path throughout the entire Web site, but also improves the usability of the Web site, making it easier for a user to navigate through the entire site.

Dickson and Marshall (2004:28–36) also concluded, through several interviews, that getting a site/page referenced by other third-party Web sites is one of the biggest boosters to the credibility of the content of a site. It was also mentioned that if other Web sites reference a site, especially if they are in the same subject community, it means that the content is likely to be good. This is because peer sites rate the content high enough to link it (Dickson and Marshall 2004:28–36).

An online newsletter can increase the chances of users coming back after the initial visit (Van Steenderen 2001). Allowing users to subscribe/unsubscribe directly from a Web site is a low cost alternative to keep users up to date on the latest offerings from the site. Another strategy to gain Web site popularity is to update it on a regular basis.

Another way to improve the perception about the size of a company, is to provide different e-mail contacts e.g. info@SA-Cycling.com, orders@SA-Cycling.com, david@SA-Cycling.com and point them all to the same inbox. This will create the perception that the customer is dealing with a large company, and is suggested in cases where a site is designed to operate as an e-commerce business.

4 Results

The only major perceived difference between the old site (see Appendix A, Figure 2) and the new improved site (see Appendix A, Figure 3) is the menu bar in the left margin. Before the changes were implemented, the menu was created in JavaScript. It looked good, worked well and was easy to use. However, most crawlers were unable to read the code used to write the menu. Therefore the menu was changed to a column with text in multiple rows linking to appropriate pages. Each of the menu options can now easily be read by a crawler, allowing the menu options to be listed as key words. More importantly, crawlers could now also follow the embedded hyperlinks to other pages on the same site.

The second noticeable change is the heading. It now consists of ordinary text, starts with capital letters and is written in bold. The text is in larger type and bold, and it is expected that the new heading would receive a much higher ranking value by search engine algorithms.

The third noticeable change is the title of the Web page. It has been changed to a more, key word-dense sentence that accurately describes the content of the particular Web page, rather than just the site name and address. It is expected that this should also highlight some key words associated with the Web site.

The fourth and final noticeable change concerns the frames. The removal of the frames should eliminate the risk that some search engine crawlers would not be able to find key word rich Web pages on the site. It also eliminates the chance of a crawler indexing the outside of the frame as a distinct page, and in the process miss the actual page containing the body text.
The ALT tags (text describing the graphics) are now visible when graphics are turned off. This improvement is expected to allow end-users, who turned the graphics off to find their way around the site, to read the graphic descriptions. At the same time the descriptive text should provide valuable key words for the crawlers to interpret. See Appendix B, Figure 4 for an example of the value of ALT tags.

The site map (Figure 1) that has been added now provides several links to the rest of the site. This allows most search engine crawlers to read every page on the site. Not only should it increase the visibility to the crawler but it should also increase the usability of the Web page to human visitors. Figure 1 shows the design of the implemented site map.

**Figure 1 Site map**

- **Services**
  - Trade Zone
  - Directory
  - Email Lists
  - Stolen Bikes
- **Journals**
- **Books**
- **Products**
  - Standard Products
  - New Products
  - Feature Products
  - Specials
- **Support Plans**
- **Customer Support Login**
- **News**
  - Overview
  - In the News
  - Press Releases
  - R.I.P.
  - History
- **Clue**
- **Touring**
- **Conferences and Events**
- **Company**
  - Overview
  - Executive Team
  - Subscribe
  - Contact Us

**5 Conclusion**

In this article, a proposal was made on how to use the latest known good practice prescriptions for improving Web site visibility. Should these changes be implemented, the Web site owner can expect the following:

- Higher Web site visibility;
- increased crawler friendliness without sacrificing human friendliness;
- increased Web site usability to disabled users by implementing text that describes the graphics;
• increased usability made possible by the site map; and
• ultimately, higher Web site traffic and increased income.

Meta, TITLE and ALT tags are important and should be created in such a way that they will serve as an online marketing campaign for the Web site. A lot of time should be spent on these tags, because well-written and effective tags will not only improve the user's experience of interacting with the Web site, but also increase its rankings.

A Web site should always be written as simple as possible, for the sake of being easily indexed by search engine crawlers. Most search engines like text-based, content rich pages. They find it difficult to index Flash or JavaScript, they cannot see graphics and they also misinterpret framed sites.

Users want to visit a site and easily find what they are looking for in the shortest time possible. They do not want to have to sit through lengthy Flash movies. They want simple HTML text links with clear navigation throughout a site.

Information provided in the simplest way to search engines and human users alike should result in both search engine and user satisfaction, thus increasing traffic which brings in new business.

Future research could include the following:

• Implementing the newly designed SA-Cycling Web site, submitting it to search engines and monitoring its ranking; and
• analysing a chosen Web site's traffic with WebPosition Gold's Traffic Analyzer which tracks visitors using an invisible counter on the Web site.

6 Acknowledgements

The authors wish to thank the National Research Foundation (NRF) for partial funding, the Cape Peninsula University of Technology for the provision of a research assistant, and David Cowie for providing his Web site to make this research possible.

7 References


Galon, D. 1999. The savvy way to successful Website promotion: Secrets of successful
Websites, attracting on-line traffic. The most up to date guide to top positioning on search engines. Canada: Trafford Publishing.


Weideman, M. 2004b. Empirical evaluation of one of the relationships between the user, search engines, metadata and Web sites in three-letter .com
Appendix A

Figure 2 Site before changes were implemented

![Site before changes were implemented](image)

Figure 3 Site after changes were implemented

![Site after changes were implemented](image)
Appendix B

Figure 4 Web page loaded without graphics
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.