Website navigation architectures and their effect on website visibility - a literature survey.

MELIUS WEIDEMAN
and
MONGEZI NGINDANA
Faculty of Business Informatics, Cape Technikon

Search engines hold a promise of delivering relevant and useful information to the human user. The primary objective of this research project is to compare and report on different types of navigation schemes, their advantages and disadvantages, and the impact they have on the visibility of a webpage to a search engine crawler.

The method employed was to review relevant literature, compare the advantages and disadvantages of navigation architectures and to reach a conclusion. It was found that a number of options are offered to the designer of a website, including text-based links, navigation buttons, image maps, JavaScript, Flash elements, hidden menus and frames. All of these elements appeared to have a question mark over their positive contribution to the visibility of a webpage.

The primary conclusion reached is that navigation architecture used on a website does impact its visibility to a search engine crawler. The webpage designer should exercise care in choosing a navigation scheme. One option is to duplicate navigation schemes to please both human and crawler visitors, which could add to clutter and information overload. Finally, some areas for further research are identified.

Categories and Subject Descriptors: H. [**Information Systems**]: H.3 Information storage and retrieval; H.3.3 Information Search and Retrieval.
General Terms: Design, Human factors, Performance.
Additional Key Words and Phrases: navigation, visibility, search engine, crawler.

_____

1. INTRODUCTION

The navigation scheme of any website plays an important part in the user experience of that website, and clutter is a major obstacle in its design [Nielsen 2000:221]. There appears to be two opposing requirements for designing this scheme: one to please the human visitor and one to please the search engine crawler visitor. Some authors claim that many developers design web pages that are created entirely in Flash or with graphic images to preserve the aesthetics of the colours, uncommon typefaces and movement to satisfy users, forgetting that these pages are not search engine friendly [Lide and David 1999]. Others claim the opposite: the focus should be on the crawler, to ensure ease of indexing when these programs start their visitations [Weideman 2003].

Thurow [2002] reports that, in order to get best search engine visibility, web designers should follow the 'Five Basic Rules of Web Design'. These rules stipulate that websites should be:
- – easy to read,
- – easy to navigate,
- – easy to find,
- – consistent in layout and
- – consistent in design.

By following these rules, the designer ensures that the website satisfies the target audience, as well as search engine crawler programs.

This research project compares and evaluates different navigation schemes and their impact on search engine visibility.

Author Addresses:
M. Weideman, Faculty of Business Informatics, Cape Technikon, PO Box 652, Cape Town 8000 SOUTH AFRICA meliusw@yahoo.com
M. Ngindana, Faculty of Business Informatics, Cape Technikon, PO Box 652, Cape Town 8000 SOUTH AFRICA mongeziw@ctech.ac.za

1.1 Background and Problem statement

It has been reported that HTML coding or scripting can prevent a spider from crawling web pages. Some examples of site navigation architecture that can be problematic include the following:

- poor HTML coding on all navigation schemes,
- the use of image maps (many search engines do not follow the links inside image maps),
- frames (some search engines can follow the links on a framed site, but the manner in which pages are displayed in search results is not ideal), and
- JavaScript (the major search engines do not follow many of the links, including mouseovers/rollovers, arrays, and navigation menus, embedded inside JavaScript) [Thurow 2002].

Whalen [2003] identified a number of issues as being problematic in terms of search engine visibility:

- Flash movies,
- DHTML,
- JavaScript rollovers,
- style sheets,
- images and
- animation.

Although these features can help visitors navigate with greater ease and make the website easier to read, they can interfere with search engine visibility [Thelwall 2000]. Search engine crawlers look for text on a web page, and web sites which are text-poor, present these programs with very little to index.

The problem identified at this stage is the apparent contrast in requirements for human friendly as opposed to crawler friendly webpages.

2. LITERATURE REVIEW

One of the most important components of a visible website is the site architecture [Thelwall 2000]. Site architecture includes a web site's navigation scheme and layout of elements on pages.

Alexander [2002] claims that site architecture is very important, because the search engine spiders must be able to find and record the keyword-rich text on your web pages. Without effective site architecture, web site marketers must pay for search engine visibility through paid inclusion or pay-per-click programs. A discussion of a variety of navigation architecture options follows.

2.1 Text-based links

Several authors [Ollins 2000; Thurow 2002; Murata 2003] concur that all search engines spiders successfully interpret text links because they can record the text in and around the link and follow these links from one to the next webpage. Usability experts also propose the use of text links because they provide the target audience with valuable information about visited and unvisited pages [Nielsen 2000].

However, Tansley [2002] claims that too many text links on a webpage can dilute keyword density and interfere with keyword prominence. Human browsers also do not find webpages containing only text easy to read [Weideman 2004]. Keyword prominence can be retained by the use of bread-crumb links [Aguillo 2000]. This author also states that if there are more words in the text links of a given webpage than in the main body content, the designer should consider using graphic images as part of the navigation scheme.

A navigation scheme based on text-links might seem like an ideal solution because all the search engines prefer this type of link. A web page full of text-link tends to download much faster than a web page full of graphic images. The question is whether the lack of graphic images as a user-friendly method of linking will not present the user with too much text too read, resulting in a human unfriendly navigation scheme?

2.2 Navigation Buttons

A navigation button is a graphic image, generally in the GIF or a JPEG format, that links to a single URL. Many designers choose to place navigation buttons at the top and along the right side of the screen [Greenough et al 2001].

Navigation buttons give website visitors a visual representation of how to navigate the website. They are visually appealing and can easily draw attention to important parts of the website. These buttons should always contain alternative text in case the target audience cannot view the graphic images.

Search engines can follow the link coupled to the navigation button, as long as the navigation button does not contain JavaScript within the anchor tag. JavaScript can pose problems to search engine crawlers.

The question now remains if the extra download time and storage space required for these images weigh up against the advantages they offer?

2.3   Image Maps

An image map is single graphic image that enables users to access different webpages by clicking different areas of the image. An example is a graphic image of a country map, where different geographical areas provide links to weather forecasts for each individual area [Weathersa 2004].

According to Whalen [2003], many search engines do not follow the links inside an image map because of the possibility of spam. Furthermore, search engine spiders assess the main body text on a web page with more ease if the navigation scheme is a single graphic image rather than multiple ones. However, a single graphic image downloads much more quickly than multiple graphic images. Once again it is not clear which option provides the best value for money.

2.4   JavaScript

JavaScript is a programming language that enables web site designers to add flair and interactivity to their web sites. Some of the most common uses for JavaScript include:
  −   rollovers/mouseovers,
  −   pop-up windows,
  −   form validation, and
  −   drop-down navigation menus.

A JavaScript navigation scheme can greatly decrease the ease with which a crawler can index a webpage [Murata 2003]. Currently, most search engines do not follow the links embedded inside JavaScript code.

One author has proven that sites without JavaScript in the navigation scheme consistently rank higher in search engine results than the sites with JavaScript in the navigation scheme [Rao 2000]. This author further reports that webpages using JavaScript in its navigation can rank well in the search engines as long as a spider-friendly navigation scheme (text links) is also present on the page.

This finding seems to confirm that text-based links increase the visibility of a website, as mentioned earlier.

2.5   Animation/Flash elements

Flash animations and other variations provide an effective way of creating webpages with visual flair. Web designers often use this technology since it provides a way to deliver vector images, Flash movies, Splash pages and Flash animations.

According to Munafo [2002], Flash design has benefits for both web designers and site visitors. However, Flash navigation schemes contain very little text for the search engine crawler to index. Thelwall [2000] confirms that a splash page contains little or no text for search engines spiders to follow. Most splash pages contain a redirect instruction after the Flash animation has terminated - most search engines consider redirections to be spam.

Thurow [2002] claims:
'If you want to keep the flair and ambiance of Flash movie, place the text and links below the fold'.

2.6   Drop-down/Pull down Menus

The main advantage of using drop-down menus in a navigation scheme is a saving in screen real state. Drop-down menus do not initially take as much space as a series of navigation buttons or text links [Whalen 2003]. Nielsen claims that screen real estate is valuable, and only a certain fraction should be sacrificed to the navigation scheme - real content should occupy the most prominent place [Nielsen 2000].

Munafo [2002] found that drop-down menus are generally not search engine friendly because they require either JavaScript or a CGI program to work. If they are used there should be an alternative form of navigation for search engines to follow.

2.7   Dynamically Generated URL's

Search engines marketers found that not only can search engine not follow the links inside dynamically generated webpages; they also have trouble in interpreting the URL used to retrieve web pages [Aguillo 2000]. According to Ollins [2002], some examples of symbols that are not search engine friendly, that acts as indicators of dynamically delivered content includes: ?, &, $, and %. Dynamically generated webpages have URL's which contain some of these characters.

Thurow [2002] concluded that some dynamically generated URL's can trap search engine crawlers and cause them to crash. This author further concludes that search engines update their indices every four to eight weeks.

If they include dynamically generated URLs in their search results, the content might change between the time they recorded the URL and the timer the URL appears in the search results. The search results obtained by users will therefore regularly be out of date.

Sullivan [2001] lists a number of workarounds that could make dynamic sites more search engine friendly. These include:
   – minimize the number of parameters in the URL,
   – add static information,
   – add HTML pages that the search engines can index,
   – modify the stop characters in the URL,
   – utilise pay-for-inclusion (PFI) programs, and
   – participate in pay-for-placement (PFP) advertising.

2.8   Frames and Navigation

The use of frames allow website designers to provide the user with a more friendly navigation scheme, for example by anchoring one section of the screen (containing navigation elements) while the user scrolls up or down in another.

Search engine crawlers will index text and follow links in for example a navigation frame, but the initial frameset code does not present these crawlers with keyword-rich text to index and set links to follow [Murata 2003].

This author concluded that if a site uses a frames-based design, it should always include keyword-rich content and links to the most important pages within the rest of the website, so the crawlers can record the most important text.

3.   CONCLUSION AND RECOMMENDATIONS

From the literature the reviewed, it is evident that website designers should exercise extreme care in the choice and application of a navigation scheme. The main problem identified is that the requirements for a human friendly scheme are in direct opposition to those of a crawler friendly scheme. Visual clutter, especially in the navigation scheme, has a negative effect on the human's experience of a webpage, and of the website in general.

A text-based link navigation scheme might seem like an ideal solution because all the search engines prefer this type of link. A web page full of text-link tends to download much faster than a web page full of graphic images. The use of text links as main navigation scheme can satisfy both the search engines and the target audience. Building a site that appeals to both directories and the target audience is very important for maximum search engine visibility.

Navigation buttons should be used with great care. It is suggested that images which are both simple and small (in terms of disk storage, and therefore download time) are used. Furthermore, if the same navigation image set is used on all webpages of the website, download time will be reduced since each image will only be downloaded once. Finally, all images should have alternative text, and contain no JavaScript in the anchor tags.

No clarity could be found on the use of image maps in websites. While enhancing the human experience and providing a relatively easy navigational guide, download times and crawler hostility are serious problems hampering their usefulness. It is recommended that if a designer chooses to use an image map as part of the website's navigation scheme, text links or navigation buttons should be included elsewhere on the page.

JavaScript seems to offer a host of features to the designer, which enables him or her to make a solid contribution towards a visually appealing webpage. However, it is clear from the literature that JavaScript and search engine crawlers do not mix well. The suggestion of adding a text-based navigation linking scheme on top of a JavaScript based scheme, seems to defeat the objective of simple webpage layout - two navigation schemes for every webpage appears to be excessive. Not only does it add more visual and other elements for the user to interpret, but it also generates duplication.

The negative effect of the presence of a Flash movie can be cancelled by adding keyword-rich text and important links from a splash page without losing the impact of the Flash movie. However, it is doubtful whether the visual advantage of Flash elements is worth the potential user irritation of these interventions.

The use of drop-down and pull-down menus could lead to problems with website visibility. Once again, if an alternative menu system is provided in parallel to for example a drop-down system (as suggested in the literature), clutter and information duplication could result.

Dynamically generated URL's pose problems to search engine crawlers. A number of solutions are proposed in the literature, but once again some of them involve duplication which could add to clutter and information overload to the website user. Others have financial implications, which could require more resources than what the owners may be prepared to invest.

Frames-based webpages present difficulties to crawlers, since each frame in essence represent a different URL. Some solutions are proposed, which involve duplication of elements to satisfy data-hungry crawlers.

In final conclusion, it appears as if most of the solutions to the problems generated by some navigational schemes proposed in the literature could create new problems of their own. Duplication of existing information, creating extra clutter and adding to download delays are some of these problems. Potential future research in this area could look at this duplication element, the irritation factor of Flash components (both viewing and installing of prerequisite programs) and download delays.

4.   ACKNOWLEDGEMENT

5.   REFERENCES

AGUILLO, I. 2000. A New generation of tools for search, recovery and quality evaluation of World Wide Web medical resources. *Journal of Management in Medicine* 14(3), 240-246.

ALEXANDER, J. 2002. *How to Choose Keywords Before They Skyrocket in Popularity Search Engines Workshop*. http://www.searchenginesworkshops.com/articles/response.html.

GREENOUGH, R., FAKUN, D., KAY, J. 2001. Development of a digital manual for a manufacturing system – a case study. *Integrated Manufacturing Systems* 12(6), 387-393.

LIDE, R. and DAVID, M. 1999. Provant launches new e-commerce Web site, grows operating income 91.6% in 1Q. *Lifelong Learning Market Report* 4(11), 3-4.

MURATA, T. 2003. Visualizing the structure of web communities based on data acquired from a search engine. IEEE Transactions on Industrial Electronics 50(5), 860-866.

MUNAFO, R. 2002. Programming Languages: The Internet's current opinion at MROB. http://home.earthlink.net/`mrob/pub/lang_srom.html.

NIELSEN, J. 2000. Designing Web Usability. New Riders, IN, USA.

OLLINS, D. 2002. Professional Web Design – some good ideas for web design. http://www.webdevelopersjournal.com/articles/Professional.html.

RAO, S.S. 2000. E-commerce: The medium is the mart. New Library World 101(1154), 53-59.

SULLIVAN, D. 2001. *How Search Engines Work: How To Really Profit From SEO*. http://www.searchenginesbook.com/presskit.html.

TANSLEY, D. 2001. *Create Dynamic Web Pages Using PHP and MySQL*. Addison-Wesley, New York.

THELWALL, M. 2000. Effective web sites for small and medium-sized enterprises. *Journal of Small Business and Enterprise Development* 7(2), 28-35.

THUROW, S. 2002. *Search Engine Visibility Kit*. http://www.searchenginesbook.com/presskit.html.

WEATHERSA. 2004. *South African Weather Service*. http://www.weathersa.co.za/.

WEIDEMAN, M. 2003. Payment for increasing website exposure in search engine results - technical and ethical issues. In *Proceedings of The 5th annual Conference on WWW Applications*, Durban, South Africa, September 2003, CD-ROM.

WEIDEMAN, M. 2004. Ethical issues on content distribution to digital consumers via paid placements as opposed to website visibility in search engine results. In *Proceedings of ETHICOMP 2004*, Syros, Greece, April 2004, Eds. Terrell Ward Bynum, Nancy Pouloudi, Simon Rogerson, Thomas Spyro, University of the Aegean, Syros Volume 2, 904-915.

WHALEN, J. 2003. *Search Engines Marketing vs Search Engines Optimization*. http://www.searchengineguide.com/whalen/2003/0612-jwl.html.