



silkmoth
digital marketing

Getting more than your “fair share”

Utilising search engines and user behaviour to attract a greater number of useful inquiries from your web site

a white paper by **silkmoth** ltd

21 March 2005

document ref: NS101-2

Read this white paper if you want to attract more business through using the web. You may already have a web site, or be considering one for the first time. The reader is not required to have a deep understanding of web technology, and so the document is useful to managers in business as well as IT professionals.

Summary

Maximising the potential of the web is important for any business. Fundamentally, a web site should be designed to attract more visitors than competitors' web sites, with a higher than average likelihood of those visitors converting into customers.

The web is considered by many to be too large, too complicated and too dominated by "geeks" to be a serious business channel that can be analysed, measured and subject to the prediction and planning that would apply to other routes to market. Many companies invest in the development of a web-site without a clear understanding of how the web works and how potential customers use the web to satisfy their requirements.

The web continues to develop; so what may have worked well in web site design a year or so ago is not likely to be the optimal solution today. Search engines, which are the principal method by which new visitors arrive at a site, are constantly becoming more sophisticated in how they present relevant results. And the behaviour of users continues to develop as they become more and more adept at using the available tools.

This white paper puts the current technologies into context with these issues, in order to explain the reasons for failure and the best approach to assure success. A new approach developed by silkmoth ltd, called netSITE™, is described and compared with existing methods.

Types of current sites

There are two basic types of web site technology available to most businesses.

Static site

The first is the static web site which is usually quite small in terms of the number of pages it contains and often the content does not change very often. A static site would typically be fewer than 50 pages.

We can think of static sites as typically being like a corporate brochure. There may be some means for regular content management but the sites are usually small and relatively static in the content they deliver.

Database driven sites

The second is a site that is dynamically created based on content that is held in a database. This technology (because it tends to be more expensive initially, although cheaper per page thereafter) is usually used for larger sites with perhaps a range of different products or services to offer.

Database driven sites are used to speed up the building of sites that deliver potentially a large number of different pages. They may contain details of a catalogue of products. Sites built using this technology often follow through on visitors to complete a sale. The potential for sales revenue is often the justification for the higher initial cost.

How search engines work

Currently the main search engine is Google, which has about 70% of the market and is generally considered to be the best tool available to find relevant content on the Internet. We therefore focus on this search engine for this discussion.

Google uses two principles to determine which results it returns to its users:

- Page rank (importance)
- Textual content (relevance)

Page rank (importance)

Google continually assesses all the web pages that it has in its index (currently about 8 billion) to calculate the relative importance of each page, independent of the content that is on the page. The way that it does this is by an iterative approach using the pages on other web sites that link to the target page. The principle is simple and effective and produces results that it is difficult to "spam" (influence unfairly).

Page rank is expressed as a number from 0 to 10 and is based on a logarithmic scale, the details of which are a closely guarded Google secret. However the basic principle of the method is well known and is broadly as follows. If a web page is linked to from another page with a page rank of 5, for example, the target page receives a part of that referrer's page rank. If it is the only page referred to by that page it gets a larger share: if it is one of many referred, it gets a smaller share. If the link comes from a site with a large number of outgoing links (a link farm) this can actively damage the target site's page rank.

Textual content (relevance)

Google stores a database of words (which is limited to about 15 million - in English) against which it holds a list of web pages. When it indexes a web page it “reads” the content on the page in order to maintain this list. It also uses the page description and page title and any “alt-tags” on images (the text that pops up when you move the mouse over an image on a web page). Google doesn’t use keyword alt-tags although other search engines do. So, in essence, Google uses all of the content that will be visible to the end user to build a list of words that tell it what the web page is about.

When a visitor to Google searches for a phrase, Google first builds a list of web pages that contain each of the words in that phrase and then applies a number of algorithms to that list based on things like the proximity of the words to each other on the page and the presence of similar words.

How Google produces its results

Google derives a list of potentially relevant pages (from cross referencing its indexed pages and its word database) and it can then present a list to the end users based on that list and the page rank of the pages in that list.

We don’t believe that Google currently interprets meaning or applies any intelligence to the content as such. Its sole objective is to try to present pages that are likely to provide a “good” result for its users. It will therefore try to exclude pages that are simply lists of other pages (meta searches) or pages that have been built solely to achieve top search engine positioning without any underlying value.

How potential customers use search engines

Most Internet users will start searching using a one or two word search phrase. Typically such a search will produce a “potential” list of hundreds of thousands or even millions of web pages. The ones that are presented at the top of the list will usually be those with a high page rank and this may not reflect what the user is really looking for.

As an example, if we consider a typical visitor to one of our sites. The table below shows the number of results that Google returns in response to the query (in its database of UK resident sites) and an estimate of the number of searches made for that phrase.

Search phrase	Number of pages (UK) returned by Google	Number of UK searches each month
compensation	3,100,000	138,582
compensation claim	761,000	40,332
accident compensation claim	207,000	9,471
trip accident compensation claim	41,300	2,047
trip paving slab accident compensation claim	3,120	152
trip paving slab macclesfield accident compensation claim	42	1

As the search becomes more specific, the number of possible matches reduces and it becomes “easier” to get to the top of the results list. The number of people searching becomes fewer also, but each one is much more highly qualified and likely to convert to an enquiry.

Also, the total number of people using very detailed searches is actually far more than making general enquiries.

Consequence of search engine technology and user behaviour

Static site

Most static sites (because of their limited volume of content) are attempting to capture visitors from fairly generalised searches. Unless the site has a high page rank (which is difficult to achieve), the site is unlikely to appear within the top 20 results in Google. It is therefore unlikely to achieve any significant traffic.

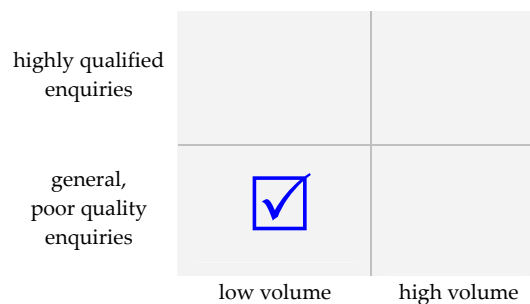


Figure 1: Static Site

Database driven site

If a database driven site is delivering content that changes regularly there is a high probability that the results that appear in Google will be out of date.

Even if the content is not changing regularly, search engines will not successfully index most of the content because usually selections in database driven sites are selected by means of drop down lists or “radio buttons” and search engines will not be able to run these searches in order to generate the pages to index.

The consequence of this is that a database driven site can produce a mix of a few highly qualified enquiries (where the search engine happened to index an appropriate page recently enough for it to satisfy the web searcher), or a larger number of poor quality enquiries (where the search engine cannot present up-to-date and specific enough results).

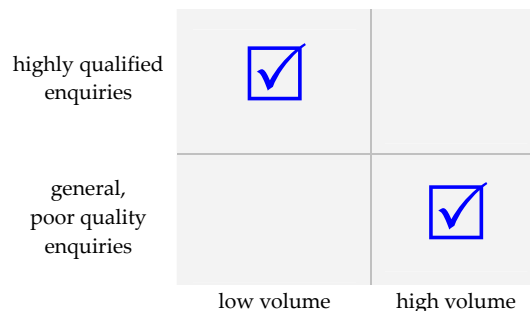


Figure 2: Database Driven Site

A new alternative: netSITE™

Silkmoth's netSITE™ uses patent pending technology, combined with an analysis of actual search engine traffic generated for particular search phrases, to build sites that generate high volumes of well qualified enquiries at a lower cost than the other options. These web sites can capture visitors making highly detailed search queries which are made up of, typically, three or more keywords.

We use our patent pending technology to build a multi-dimensional matrix of key words, phrases or concepts and to automatically generate a web site. They are also optimised for the way that search engines use "robots" to index web sites.

netSITE™ solves the problems experienced by web sites designed in the existing manner, by automatically generating site content that is static, yet constantly refreshed, so that the search engines record, index and deliver "current" content. This would be impractical if it were not for the automation inherent in netSITE's creation and management of a web site.

Creating highly targeted, wide ranging sites, causing the search engines to visit and index. By actively monitoring, reporting and analysing the search engine and user traffic we glean maximum information and value from the site.

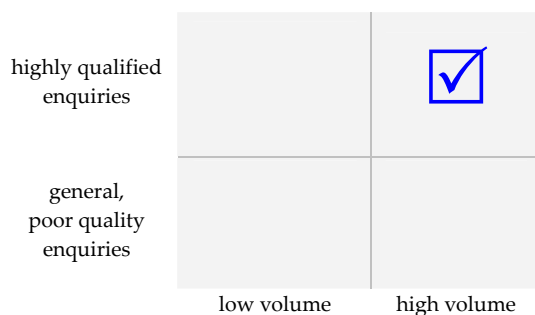


Figure 3: netSITE

Conclusion

Selecting the right technological approach for the creation and maintenance of a web site is essential if the web site is going to produce optimal results for the site's owner. This discussion shows that the behaviour of both search engines and web users has to be understood properly if a sound decision on the right approach is to be made.